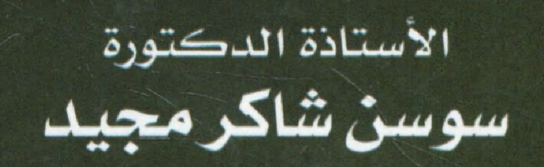
أسس بناء الاختبارات والمقاييس النفسية والتربوية



الأستاذة الدكتورة
سوسن شاكر مجيد

مركز ديبونو لتعليم التفكير
De Bono Center for Teaching Thinking


مركز ديبونو لتعليم التفكير
De Bono Center for Teaching Thinking
مركز ديبونو لتعليم التفكير
تفكيرنا... مستقبلنا

أسس بناء الاختبارات والمقاييس النفسية والتربوية

إعداد

أ. د سوسن شاكر مجيد

جامعة بغداد/ كلية التربية

الناشر

مركز ديبونو لتعليم التفكير

رقم التصنيف: 370.15

المؤلف ومن هو في حكمه: أ.د سوسن شاكر مجيد

عنــوان الكتــاب: أسس بناء الاختبارات والمقاييس النفسية والتربوية

رقم الإيـــداع: (2006/12/3299)

الترقيم الدولي: ISBN: 978-9957-454-25-0

الموضوع الرئيسي: الاختبارات التربوية/ علم النفس التربوي/ التعلم

* تم إعداد بيانات الفهرسة والتصنيف الأولية من قبل دائرة المكتبة الوطنية

الطبعة الثالثة

2014م

مركز ديبونو لتعليم التفكير

عضو اتحاد الناشرين الأردنيين

عضو اتحاد الناشرين العرب

يطلب هذا الكتاب مباشرة من مركز ديبونو لتعليم التفكير

عمان- شارع الملكة رانيا- مجمع العيد التجاري - مبنى 320- ط 4

هاتف: 962-6-5337003 / 962-6-5337029

فاكس: 962-6-5337007

ص. ب: 831 الجبيهة 11941 المملكة الأردنية الهاشمية

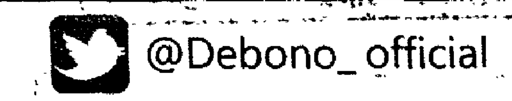
E-mail: info@debono.edu.jo

www.debono.edu.jo





debonotrainingcenter



المحتويات

مقدمة

تعد عملية القياس والتقويم من العمليات المهمة والقديمة قدم تاريخ الإنسان نفسه منذ أن وجد على وجه الكرة الأرضية. فقد استخدم الإنسان القديم الطرق البدائية الساذجة البسيطة في عملية القياس والتقويم. فعندما كان يحاول اصطياد حيوان ما فإنه كان عليه أن يقدر المسافات وتحديد حجم الصخرة التي يقذفه بها لاصطياده. وعندما كان ينظر إلى جماعته من حوله فإنه يعرف أن زميله الضخم أقوى من في الجماعة وان الآخر أسرعهم في العدو. كل هذه الأمور الحياتية وغيرها تحتوي على عمليات القياس والتقويم.

واستخدمت الصين القديمة نظام للامتحانات التحريرية قبل أكثر من ثلاثة الآلاف عام كأساس للقبول والترقية في الخدمة المدنية.

ولدى الإغريق استخدم التقويم للسيطرة على المهارات البدنية والعقلية. وعرف العرب القدامى معنى التقويم منذ القدم وتمثل ذلك بتقويم نتاجا تهم الفكرية شعراً ونثراً من خلال عقد الندوات في الأسواق كسوق عكاظ وفي مواسم الحج، ثم يقوم الخبير بإصدار أحكامه معتمدا على معايير متفق عليها فظهرت على هذا الأساس "المعلقات السبع" التي اعتبرت من القصائد القمم.

ثم جاء الإسلام الذي يعد نظاماً تربوياً شاملاً، فأصبح العرب المسلمون يقومون سلوك الإنسان بناء على مدى انطباق تعاليم الإسلام على سلوكه. وتوجد الكثير من الآيات القرآنية الكريمة التي تتضمن إشارات صريحة إلى عملية القياس والتقويم وفق معايير الإسلام مثل ﴿ هَلْ يَسْتَوِي الَّذِينَ يَعْلَمُونَ وَالَّذِينَ لَا يَعْلَمُونَ ﴾ (الزمر: 9)، ﴿ فَمَن يَعْمَلْ مِثْقَالَ ذَرَّةٍ خَيْرًا يَرَهُ ۝٧ ﴾ (الزلزلة: 7)، ﴿ وَمَن يَعْمَلْ مِثْقَالَ ذَرَّةٍ شَرًّا يَرَهُ ۝٨ ﴾ (الزلزلة: 8)

وفي الوقت الحاضر يعني النظام التربوي والتعليمي بعملية القياس والتقويم، إذ أن نتائج التقويم توجه المربين نحو اختيار أهداف معينة وتحقيقها فهي عملية

ضرورية للمعلم والمتعلم. ضرورية للمتعلم لأنه بتقويم أدائه يستطيع أن يتعرف على مستواه ويقارنه بما بذله من جهد، الأمر الذي يحفزه على بذل المزيد من المثابرة والتحصيل لتحقيق طموحه. وضروري للمعلم أن يتمكن من خلاله القيام بعملية تقويم تحصيل تلامذته ومستوياتهم العقلية والنفسية ويساعدهم على اكتشاف أنفسهم من جهة وعلى تحقيق أهداف التربية من جهة أخرى. كما أنها ضرورية للقائمين على أمر التربية والتعليم لأنه يساعدهم على وضع الاختبارات النفسية والتربوية بما تتلاءم مع طبيعة البرامج التدريسية والتربوية والإرشادية. وأن طالب الدراسات العليا يحتاج أيضاً خلال سنوات دراسته إلى الإحاطة بصورة معمقة بموضوعات ومفاهيم القياس والتقويم وأساليب بناء الاختبارات والمقاييس النفسية والتربوية.

لذا فإن الكتاب الحالي يهدف إلى تقديم الخطوات الأساسية عن كيفية بناء الاختبارات والمقاييس النفسية والتربوية لاستخدامها بعد إنهائه دراسته وتخصصه. إن هذا الكتاب تنحصر موضوعاته كليا في عملية القياس والتقويم في المدرسة والأسس المعتمدة في بناء الاختبارات والمقاييس النفسية والتربوية، وان تطرقت إلى موضوعات أخرى خارج عن الموضوع.

يقع الكتاب في (12) فصلا إضافة إلى المقدمة والمراجع. وقد تناولت في **الفصل الأول**، مفهوم القياس والاختبار والتقويم، نبذة تاريخية، ومفهوم القياس، مستويات القياس، مفهوم الاختبار، مفهوم التقويم، مفهوم التقويم والقياس والعلاقة بينهما، استخدام الاختبارات، أنواع الاختبارات، العوامل التي تؤثر في القياس، مبادئ عامة ينبغي مراعاتها في التقويم.

أما **الفصل الثاني** فقد تناولت الأهداف التعليمية، وتعريف الهدف السلوكي، وأهمية تحديد الأهداف، وهل الهدف ضرورة ملحة في العملية التربوية؟، إيجابيات الأهداف في العملية التربوية، تصنيف بلوم للأهداف التعليمية، تصنيف زايس، تصنيف انيتاهارو، صياغة الأهداف السلوكية، اعتبارات ينبغي مراعاتها عند الأخذ بالأهداف السلوكية، دور الأهداف السلوكية في العملية التعليمية.

أما **الفصل الثالث** فقد تناول خطوات بناء الاختبار أو المقياس، خطة تصميم الاختبار، تحديد الهدف أو الأهداف الرئيسة للاختبار، تقرير محك أو معيار الدرجة، ترجمة المفاهيم والأهداف إلى خصائص محددة، تصميم فقرات مناسبة تعبر عن هذه الخصائص، اختبار مستوى الصدق والصعوبة للفقرات، فعالية المشتتات، إعداد الاختبار للاستخدام تقنين الاختبار.

أما **الفصل الرابع** فقد تناول مفهوم الصدق، تعريفه، أهميته، أنواعه، المحتوى، الظاهري (السطحي)، المفهوم، التجريبي، المحك (التنبؤي، التلازمي)، البنائي، العاملي، ومقارنة بين أهم أنواع الصدق، الطرق المختلفة لحساب الصدق، تصحيح معامل صدق الاختبار، العوامل المؤثرة في صدق الاختبار.

أما **الفصل الخامس** فقد تناول مفهوم الثبات، تعريفه، أنواع معاملات الثبات، نظرية الثبات، العوامل المؤثرة في ثبات الاختبار، أساليب حساب معامل الثبات، أعادة الاختبار، الصور المتكافئة، التجزئة النصفية، تحليل التباين، ثبات المصححين.

أما **الفصل السادس** فقد تناول المعايير، الحاجة إلى المعايير، الخصائص، المعايير والتقنين، أغراض المعايير، الدرجة المعيارية.

أما **الفصل السابع** فقد تناول الاختبارات المحكية المرجع والاختبارات المعيارية المرجع والاختبارات التشخيصية، أوجه الاختلاف بين الاختبارات معيارية المرجع ومحكية المرجع، تصنيفات الاختبارات محكية المرجع، خطوات بناء الاختبارات محكية المرجع (الجانب المعرفي)، بناء الاختبارات محكية المرجع لقياس الجانب الوجداني، طرق تقدير صدق الاختبارات محكية المرجع، طرق تقدير ثبات الاختبارات محكية المرجع، مسميات وتعريف درجة القطع، طرق تحديد درجة القطع، أسس تقويم طرق تحديد درجة القطع، الاختبارات التشخيصية، خطوات بناء الاختبارات التشخيصية، أنماط التعليم ومستوياته.

أما **الفصل الثامن** فقد تناول الاختبارات التحصيلية، الاختبارات التحصيلية

المدرسية، الاختبارات المقالية، الاختبارات التحصيلية الموضوعية، الاختبارات قصيرة الإجابة، فقرات ملء الفراغات، فقرات الاختيار من متعدد، فقرات المزاوجة، فقرات الخطأ والصواب، الاختبارات المقننة الخارجية، بناء الاختبارات التحصيلية، بنك الأسئلة، اختبارات الأداء.

أما **الفصل التاسع** فقد تناول اختبارات الذكاء، ماهية الذكاء، نبذة تاريخية عن اختبارات الذكاء، طبيعة اختبارات الذكاء، عما تخبرنا اختبارات الذكاء وبم تتنبأ؟ وحدات الذكاء، طبيعة الذكاء المقاس، ثبات اختبارات الذكاء، الوراثة، التشابه في نسب ذكاء ذوي القربى، التشابه الوراثي أو البيئي، مفهوم الموروث، ما الذي نعنيه بالموروث، نماذج من اختبارات الذكاء.

أما **الفصل العاشر** فقد تناول مقاييس الاتجاهات، مقدمة، تعريف الاتجاه، أنواع الاتجاهات، خصائص الاتجاه، وظائف الاتجاه، الفرق بين الاتجاه والعاطفة، الرأي العام، التعصب، القيم، تكوين الاتجاهات، المجالات والأغراض التي تستخدم فيها مقاييس الاتجاهات، قياس الاتجاهات، صدق مقاييس الاتجاهات، طرق قياس الاتجاهات، تحليل المحتوى، نماذج من مقاييس الاتجاهات والقيم.

أما **الفصل الحادي عشر** فقد تناول قياس الشخصية، وسائل التقرير الذاتي، المقابلة، تاريخ الحالة، مقاييس التقدير، تقدير السمات عن طريق الاستبيانات والقوائم، الاختبارات الاسقاطية، تصنيفات أخرى لقياس الشخصية، تصنيف وكنس، تصنيف وتيلا، تصنيف أنستاري، تصنيف كرونباك، تصنيف بيرت، تصنيف الزوبعي، المشكلات التي تواجه اختبارات الشخصية وقياسها.

الفصل الثاني عشر، قياس الميول، طبيعتها، علاقتها بالمفاهيم الأخرى، تكوينها، العوامل التي تؤثر في الميول، تقدير الميول وقياسها، قائمة سترونك للميول المهنية وقياسها، بيان المفاضلة المهنية لكودر، اختبار مسح الميول ليفورد وزيمومن، اختبار الميول المهنية لأحمد زكي صالح.

أسأل الله العلي القدير أن يوفقنا إلى ما فيه رضاه عني وقبول ما أكتبه من الجمهور القارئ وخاصة المتخصصين وطلبة الدراسات العليا، وأنني في جهدي هذا لا أدعي الكمال فالكمال لله وحده ... والله ولي التوفيق.

المؤلفة
الأستاذة الدكتورة سوسن شاكر مجيد
جامعة بغداد / كلية التربية ابن الهيثم

الفصل الأول

مفهوم القياس والاختبار والتقويم

يحتوي الفصل على:

- نبذة تاريخية
- مفهوم القياس
- مستويات القياس
- مفهوم الاختبار
- مفهوم التقويم
- مفهوم التقويم والقياس والعلاقة بينهما
- استخدام الاختبارات

نبذة تاريخية:

للاختبارات ماضٍ عميق فقبل أكثر من ثلاثة آلاف سنة دأب الصينيون القدامى على استعمال التقويم والقياس الموضوعي لاختيار الحكام والإداريين لمختلف مقاطعات ومدن الصين واستمرت هذه الحال لأكثر من آلفي سنة. إذ كانت تجرى اختبارات عامة للمتقدمين تكون فيها الأسماء سرية ويتولى تصحيح إجابات الطلبة أكثر من مصحح واحد، وكانت الامتحانات تجرى بشكل متتابع من القرية إلى المدينة

إلى المقاطعة يشترك فيها آلاف المتقدمين في موضوعات شاملة ومتنوعة تتضمن اللغة والحساب والشعر والتاريخ والفروسية والرماية أي أن الامتحانات كانت نظرية وعملية ويرى البعض أن تلك المدة من تاريخ الصين التي عرفت الاستقرار الإداري والسياسي كان للتقويم والقياس دور كبير فيها.

وللاختبارات دور مهم في التاريخ العربي والإسلامي خاصة في المجال التعليمي والمهني وفي المستويات كافة من الكتاتيب وحتى اختيار المعلمين والأطباء والجراحين والصيادين.

وقد صنف العرب المسلمون التعليم تبعاً للفروق الفردية لذكاء المتعلم وأدركوا بشكل واضح أهمية المفاهيم المتكاملة للتوجيه والإرشاد التربوي لوضع الشخص المناسب في المكان المناسب استناداً إلى الفروق الفردية بين الأفراد.

ويقول الغزالي عن تفاوت الذكاء ما ينص:

"إذا غلبت القوة البدنية على النفس يحتاج المتعلم إلى زيادة التعلم طول المدة، وإذ غلب نور العقل على أوصاف الحس يستغني الطالب بقليل التفكير عن كثرة العلم".

وفي التوجيه التربوي ينصح الزرنوخي بما يأتي:

"إلا يختار الطالب نوع العلم بنفسه بل يفوض من أمره إلى الأستاذ فإنه قد حصل له من التجارب في ذلك، فهو أعرف بما ينبغي لكل واحد وما يليق بطبعه. أما المدرس فعليه أن يتصفح طلابه كما يتصفح خطاباً حرره وأن لا يدع إلى حلقته إلا من كان قادراً على استيعاب ما يجري فيها".

وأستعمل العرب المسلمون الاختبارات التحصيلية على شكل امتحانات شفهية وتحريرية. وكان المعلمون والمدرسون يقومون بها بشكل دوري من حين إلى آخر ففي الكتاتيب.. إذا ما أتم الطفل مدة الدراسة وهي خمس سنوات، تعادل بشكل عام الابتدائية، يمتحن المعلم الصبي لمعرفة مدى حفظه للقران الكريم ومدى قدرته

على ضبط المحفوظات وبعض القواعد المهمة فيجتاز الامتحان باحتفال خاص يدعى بالختمة والامتحانات كانت تجرى فردية لأن التعليم فردي تبعاً لقدرة كل طفل أو صبي وكانت هناك ثلاثة تقديرات للدرجة، ممتاز ووسط وضعيف. والممتاز يعني أن الطالب يستظهر القران حفظاً من أوله إلى أخره مع ضبط الشكل والأعراب والفهم وحسن الخط والمتوسط من كان يقرأ القران نظراً في المصحف مع ضبط الشكل والهجاء، أما الضعيف فهو الذي يقرأ بدون ضبط للحروف.

وفي الحلقات العلمية كانت الاختبارات تستعمل لنقل الطالب الذي يرغب بالانتقال إلى مصاف الشيوخ أي يصبح مدرساً له حلقة به، وتجرى له عدة امتحانات على شكل جلسات يوجه له فيها العديد من الأسئلة من قبل المدرس والطلبة وتجرى المنافسة حولها حتى يقتنع به الجميع بأنه يمكن أن يكون مدرساً وآلا يعود كأحد أعضاء حلقة الدراسة ويعتبر إتقان الكتب هو المقياس لاختبار قدرة الطالب، ومتى أكمل الطالب إتقان عدد مهم من كتب الدين والتاريخ والأدب كتب له الشيخ شهادة على الورقة الأخيرة والأولى من الكتاب يشهد له بأن فلان أبن فلان أتم إتقان علوم ذلك الكتاب ويضع عليه التاريخ الهجري. ومن أقدم الشهادات القديمة الشهادة التي منحها محمد بن عبد الله الحميري في 304 هجرية إلى عامر سعيد بن عمر فيما ورد في كتاب الأستاذ.

وقد يكون العرب المسلمون أول من وضعوا الاختبارات المهنية لاختيار الرجل المناسب في المكان المناسب وعلى أساس مواصفات معينة مستجدة من طبيعة العمل أو المهنة وقد بدأ بذلك الرسول ﷺ وأوضح بأنه أختار بلالاً للأذان لأنه كان أندى صوتاً. وعهد بالقضاء إلى من عرف بالاتزان والحكمة وقيادة الجيش إلى من عرف بالشجاعة والحزم في الشدائد والحروب، وعهد بالجباية إلى من عرف أمانته وورعه.

وقد تطورت هذه الممارسات بشكل واسع حتى أصبح لكل مهنة شروط وامتحان ومتخصصون في الانتقاء يعينون من قبل الحلقة أو من ينوب عنه وعلى

سبيل المثال فقد أوكل الخليفة هارون الرشيد إلى الكساني اختيار المعلمين بعد امتحانهم، وأختار الخليفة الواثق بالله العباسي عثمان النحوي المازني عام 248 هجرية بأن يمتحن جماعة من المعلمين تقدموا لتعليم أولاده فأمتحنهم فما وجد فيهم طائلا فعرفهم.

وقد أهتم المسلمون اهتماماً كبيراً باختيار الأطباء والجراحين والكحالين والصيادلة وغيرهم من المختصين الأكثر أهمية في عصرهم.

وروى أن الخليفة المقتدر أول من شرع امتحان الأطباء وكان رئيس الأطباء هو الذي يمتحن زملاءه إذ أشترط على من يرغب في الاشتغال بمهنة الطب أن يجتاز امتحاناً ينال فيه شهادة تحدد له الأمراض التي يمكن أن يعالجها.

ويقال أن ثابت بن سنان بن قره أتصل بالمقتدر سنة 319 هجرية وأخبره أن خطأ جرى على رجل من العامة من بعض المتطببين فمات الرجل فأمر المقتدر إبراهيم بن محمد يمنع سائر المتطببين من التصرف إلا بعد أن يمتحن سنان بن ثابت بن قرة وكتب له رقعة بخطه بما يطلق له الحق في الممارسة ويكون الامتحان نظرياً في محتوى كتب الطب في ذلك الزمن وفي الغالب (كتاب مهنة الطب) لأبن ماسويه وغيره، كما يمتحن الطبيب بموضوعات علمية وتفحص الآلات التي يعتمدها في عمله.

وثمة اختبارات أخرى للجراحين تجرى من قبل رئيس الأطباء وكان الاختبار يجرى في كتاب قاجانس في الجراحات والالآتها ويطلب من الجراح معرفة ما هو معروف في تشريح الجسم وما فيه من أعضاء وعضلات وعصب وشرايين.

وكذلك المجبرون فكان يمتحنهم رئيس الأطباء ويشترط بمن يمتحن هذه المهنة أن يعرف عظام هيكل الجسم ووظيفة كل عظم.

وفي تخصص العين كان رئيس الأطباء يمتحن الأطباء في كتاب الشعر مقالات في العين لحنين بن إسحاق وكان يطلب منهم معرفة طبقات العين ورطوبتها وأمراضها وطرق تداويها والأدوية المستعملة لها.

أما الصيادلة فإن امتحانهم له خصوصية تجمع ما بين الطب والكيمياء وقد يكون الخليفة المأمون أول من عني بامتحان الصيادلة بعد أن كثرت ادعاءات المتطفلين على المهنة وكشف بنفسه زيف وكذب دعواهم بعد امتحان أجراه شخصياً فقد طلب دواءً وهمياً أسماه (شقينا) وأرسل إلى جميع الصيادلة يطلب منهم هذا الدواء وأدعى أغلبهم بتوفر الدواء لديهم وأخذوا شيئاً من حانوتهم وأرسلوا إلى المأمون بأشياء مختلفة فمنهم من أتى ببعض البذور ومنهم من أتى بقطعة من حجر ومنهم من أتى بوبر. فأمر المأمون بامتحان الصيادلة ليجيد من كان متمكناً من صنعته وليعد من كان غير ذلك.

كما أمر الخليفة المعتصم بن الرشيد الطبيب زكريا الطيفوري بأن يمتحن الصيادلة وقال له: أن ضبط هؤلاء الصيادلة عندي أولى مما تقدم فيه فأمتحنهم حتى نعرف منهم الناصح.

إن الأمثلة التي نسوقها هنا توضح لنا أن الاختبارات موضوع عرفه العرب واهتموا به في وضع الشخص المناسب في المكان، وأدركوا بوعي واضح أهمية التعليم وفق القدرات. وأن موضوع الاختبارات والمقاييس لم يبدأ تماماً بدايات مستقلة عن التراث الإنساني المتصل والذي أدى فيه العرب دوراً مهماً في تطوره هنا وهناك في اختصاصات مختلفة كانت الأساس الذي تطور منه علم نفس الفروق الفردية بشكله الحالي.

وخلال عصور التوسع الاستعماري وحينما أحتل الإنكليز الهند قبل أكثر من ثلاثة قرون أعيد اعتبار الاختبارات الانتقائية لاختبار أحكام الموظفين لإدارة مستعمراتهم وأدت الاختبارات دوراً مشابها لما أدته في مجتمع الصين.

وخلال المائة سنة الأخيرة والتي تشكل العمر التقريبي لعلم النفس توسع علم النفس الفروق الفردية واستعمالات الاختبارات بشكل واسع نتيجة أربعة عوامل هي:

1- تطور علم النفس التجريبي ودراسة ملكات وقدرات الفرد مختبرياً.
2- استعمال الاختبارات في الحروب العالمية الأخيرة ولحاجة الجيش إلى وضع الشخص المناسب في المكان المناسب.
3- استعمال الاختبارات في ميادين الصناعة نتيجة الثورة الصناعية وما نتج عنها من تطور أداري وتنظيمي.
4- تطور علم النفس الإكلينيكي (العلاجي) واستعمال الاختبارات في تشخيص الضعف والتخلف العقلي فضلاً عن تشخيص الأمراض النفسية.

وبالرغم من أن الاختبارات هي جزء من التقويم إلا أن التقويم تطور متأخراً نتيجة لنجاح حركة الاختبارات وتطورها، علاوة على أسباب أخرى فقد جاء تطور التقويم متأخراً عن الاختبارات حيث لم يتضح دوره إلا في النصف الثاني من هذا القرن وخاصة في الربع الأخير منه وأرتبط كل من التقويم والقياس بالعملية التعليمية بشكل متكامل ليصبح تخصصاً جديداً له دوره الأساس والكبير في عمليات التخطيط التربوي وفي بناء محتوى المناهج التربوية وفي طرائق التدريس إلى تشخيص صعوبات التعلم فضلاً عن قياس المعلم والمتعلم.

واستخدم الباحثون الاختبارات لأجل جمع البيانات ومعرفة الفروق الفردية في السمة المقاسة للأفراد. وقد ابتكرت مئات من هذه الوسائل ولازال يصمم غيرها باستمرار.

ويرجع الفضل في بدء حركة قياس الفروق الفردية إلى علماء الفلك ومن العجيب أن نعلم أن حركة قياس الفروق الفردية لم تبدأ على أيدي علماء النفس بل بدأت على أيدي علماء الفلك. إذ طرد أحد العلماء مساعده في مرصد كرينش لأنه تأخر عنه في رصد أحد النجوم فترة تقرب من الثانية فعزا العالم هذا إلى إهمال مساعده. وكانت الطريقة المتبعة في الرصد آنذاك هي الاعتماد على السمع والبصر وكذلك القدرة على إدراك المسافات إذ كان الراصد يضبط ساعته في وقت معين بالثانية ثم يبدأ عد دقات الساعات التي ينصت إليها أثناء تطلعه في المنظار.

ثم حدث أن قرأ العالم بيزل قصة المساعد وأهتم بمعرفة الفروق بالثواني بين تقدير أثنين من الراصدين لحركة نجم من النجوم. ثم اخذ بيزل في جمع سرعة رصد النجوم لمجموعة من الراصدين المدربين فبينت معلوماته مدى اختلاف الراصدين فيما بينهم. كما بينت مدى اختلاف الفرد بينه وبين نفسه من وقت لأخر. وتعتبر هذه المعلومات أول معلومات كمية مسجلة عن الفروق الفردية.

ولم تبدأ حركة التجريب في علم النفس إلا في عام 1876 ميلادية عندما أنشأ فونت معمله في ألمانيا إلا أن المجهودات الأولى لعلم النفس التجريبي كانت تبين مدى تأثر هذا العلم بعلم وظائف الأعضاء. إذ انحصر التجريب في قياس الإحساسات السمعية والبصرية وزمن الرجع والحركات العضلية مما يبين تجاهل علماء النفس التجريبيين الأوائل للفروق الفردية التي لم تكن في نظرهم سوى أخطاء تجريبية.

ويعزى الفضل إلى العالم فرنسيس جالتون Galton العالم البيولوجي في شق الطريق لحركة قياس الفروق الفردية على أسس سلمية متأثراً بالتطورات التي حدثت في نظرية دارون التي تؤكد على الفروق الفردية والتباين فيما بين الأفراد.

وقد أنشأ جالتون معملاً صغيراً في متحف لندن بهدف قياس قدرات الإنسان، معتقداً أن المعوقين عقلياً ينقصهم حدة الإحساس. ومقرراً أن القدرات العقلية والإدراكية قد تكون مرتبطة معاً بدرجة كبيرة. وأن إحدى تلك القدرات تكون مؤشراً للأخرى وبالتالي بدأ جالتون تقدير بعض الصفات مثل حدة الأبصار والسمع والإحساس اللوني، والحكم البصري، وزمن الرجع وقاس الأنشطة الحركية متضمنة (قوة جذب أو شد شيء ما والضغط عليه) و(قوة النفخ) أيضاً وسرعان ما أشترك بالمثل كثير من علماء النفس الآخرين في محاولة بناء اختبارات للقدرات العقلية ومنهم الفريد بينيه 1857م-1911م وهو عالم فرنسي شهير وضع أول مقياس عملي للذكاء ثم عين في عام 1904 في لجنة حكومية لدراسة مشكلات تعليم الأطفال المتخلفين حيث وضع بينيه ومعاونوه اختبار يميز الأطفال الذين يستطيعون الإفادة من التعليم المدرسي العادي من هؤلاء الذين لا يستطيعون ذلك.

وهكذا أخذت حركة قياس الفروق الفردية في الازدهار في معظم بلدان العالم الغربي وأخذت تظهر أنواع متعددة من الاختبارات والمقاييس في المجالات كافة كاختبارات الاستعدادات والتحصيل والشخصية.

مفهوم القياس:

إذا كان التقويم يعرف بتوفير معلومات موضوعية لأجل إصدار حكم فإن القياس هو التعبير الكمي لهذه المعلومات الموضوعية لغة رقمية أو حسابية بمقدار معياري (بوحدة قياس ما) تعبر عن الظاهرة المقاسة.

فالقياس هو عملية (تكميم) أو تعبير بلغة كمية أو حسابية عن صفات أو عوامل أو ظواهر لموضوعات نوعية أو معنوية أو سلوكية تتطلب إصدار حكم أو تقويم عنها.

فحينها نقول أن فلاناً ذكاؤه (جيد) فهو تقويم عام لا يعني الكثير وقد تختلف كلمة جيد من شخص إلى آخر فما هو جيد عند أحد ما، قد يكون وسطاً أو أكثر من جيد عند غيره. ولكن حينما نقول أن تلميذاً ما ذكاؤه 110 درجة من أصل 60 درجة على مقياس الذكاء، نفهم ماذا تعني كلمة جيد كمياً. ودرجات الذكاء هنا هي وحدات قياس تعبر عن صفة نوعية.

إن الأطوال والمسافات تقاس بوحدات مترية أساسها السنتمتر والمتر ... الخ وكذلك الأوزان التي أساسها الغرام والكيلو غرام وهكذا بالنسبة للزمن والحرارة فجميعها تقاس بوحدات قياسية متفق على حدودها وتكون ثابتة دائماً وتستعمل للمقارنة لتحديد الفروق الكمية بين الأشياء والعناصر أي أن القياس الطبيعي يمتاز بخاصيتين أساسيتين هما التحديد الكمي وثبات وحدات القياس.

ويختلف القياس في العلوم التربوية والنفسية عن القياس في العلوم الطبيعية

لأسباب جوهرية أهمها أن طبيعة السلوك الإنساني وطبيعة الموضوعات التربوية كالتحصيل والاستعداد الدراسي ...الخ هي موضوعات أكثر تعقيداً من الظواهر الطبيعية ولا يسهل قياسها كما في العلوم الطبيعية أما لعدم ثباتها أو لارتباطها بظواهر وعوامل يصعب عزلها بشكل مستقل لقياسها، لذلك تكون أغلب القياسات التربوية نسبية وليست مطلقة ومعرضة للخطأ أكثر مما في العلوم الطبيعية. ولا يكون القياس التربوي أو النفسي مباشراً كما في العلوم الطبيعية أنما هو استدلالي شأن أغلب الظواهر السلوكية. وهذا يتطلب أساليب متطورة ودقيقة للقياس تتوافر فيها صفات الصدق والثبات والقدرة على التمييز في المستويات الافتراضية.

وتقاس سمات السلوك كالذكاء، والقدرات الفعلية المختلفة أو الميول والاتجاهات أو أي سمة من سمات الشخصية قياساً استدلالياً غير مباشر بافتراضات لوحدات سلوكية يعبر عنها في الغالب بفقرات الاختبارات النفسية لأي من مظاهر السلوك كلما أزداد تعقد عملية القياس ارتفعت احتمالات الخطأ. وتبقى التعابير الكمية نسبية في تعبيرها فحينما يأخذ الطالب 50٪ درجة في تحصيله بدرس معين فإن 50٪ لا تعني إنه يعرف نصف المادة تماماً وكذلك لا يعني الصفر إنه لا يعرف شيئاً إنما الـ 50٪ أو الصفر هي بالنسبة إلى الأسئلة التي امتحن فيها. مثل هذه الظاهرة موجودة حتى في العلوم الطبيعية فحينها نقول أن درجة الحرارة هي صفر لا يعني ذلك إنه لا توجد حرارة أنما هي درجة حرارة لا يتأثر بها الزئبق ولو أبدلنا الزئبق بمادة أخرى لكان للصفر درجة أخرى.

وهذا يعني أن للقياس مهما كان نوعه عوامل تساعد على الخطأ أو تكون مصدراً له ولكن نسبة الخطأ في العلوم التربوية والنفسية هي أكثر مما في العلوم الطبيعية.

مستويات القياس:

وقد يأخذ القياس النفسي كما في أي علم آخر أنواعاً مختلفة أو مستويات متعددة، حيث أشار (stevens) إلى أن هناك مقاييس أسمية ومقاييس الرتبة والمسافة والنسبة، وقد طور (coombs) هذا التقسيم وحدد العلاقات التي تربط بين المستويات المختلفة وأضاف له فئات جديدة، وهي:

1- المقاييس الاسمية Nominal scales:

ويقع هذا النوع من المقاييس في أدنى مستويات القياس من حيث مدى ملاءمة الإجراءات الحسابية المعروفة ومدى تطبيقها فيه لأن الأعداد (الأرقام) تستخدم فيه لتعيين فئات ينتمي إليها الأشخاص أو الأشياء فقط، ولذلك فإن الأعداد لا تعد دالة على كميات من خصائص.

2- مقاييس الرتبة Ordinal scales:

ويستخدم هذا النوع حين يمكن تنظيم الوحدات في سلسلة تمتد بين الأدنى والأعلى في الخاصية التي نقيسها، لكننا لا نستطيع أن نحدد بدقة الفرق بين أي رتبتين ولهذا يتم في هذه المقاييس تحديد مرتبة الشيء أو مكانته في مقياس يقدم وضعاً كيفياً مثل قليل أو كثير، كبير أو صغير، الأول والثاني والأخير ...

3- مقاييس المسافة interval scales:

ويتم في هذه المقاييس وصف الشيء وصفاً كمياً في ضوء قواعد تقليدية متفق عليها حتى يمكن تحديد سعة ذلك الشيء، وتسمح هذه المقاييس بتحديد مدى بعد شيئين أو شخصين بعضهما عن بعض في الخاصية المقاسة، وأن تكون هذه المسافات في ضوء البعد عن المتوسط الحسابي بمقادير معيارية، وفي الواقع فإن أغلب المقاييس النفسية والعقلية والتحصيلية من هذا النوع، غير أن هذا النوع ليس له صفر مطلق بل صفر افتراضي.

4- مقاييس الترتيب الجزئي Partial ordinal scales:

يرى (كومبس) إنه نجد أحياناً بين فئات المقاييس الاسمية بعض العلاقات بين بعض الفئات المختلفة، ومن هذه العلاقات أن وحدات إحدى الفئات قد تكون أكبر أو أصغر من وحدات فئة أخرى رغم تكافؤ الفئات في عدد الوحدات. ولهذا يسمى هذا المستوى بمقاييس الترتيب الجزئي.

5- مقاييس الرتبة المتري Ordered metric scales:

يرى (كومبس) أن المسافات إذا كانت بين جميع وحدات الفئات ذات علاقة يصبح المقياس مرتباً ترتيباً كاملاً.

6- مقاييس النسبة Ratio scales:

وتعد هذه المقاييس أعلى مستويات القياس والتي تتميز بأن لها وحدات متساوية ولها صفر مطلق. وتصلح معها كل العمليات الحسابية وكل الطرق الإحصائية إلا أن هذا النوع من القياس لا نتعامل معه في القياس النفسي، لأننا لا نمتلك صفراً مطلقاً للخصائص الإنسانية. ويمكن استخدامه في قياس الخصائص العقلية إذا ما تم قياسها بوحدات فيزيائية كأن نقيس زمن الرجع أو التعلم بوحدات زمنية.

مفهوم الاختبار:

تعرف أنستازي الاختبار النفسي بأنه "مقياس موضوعي مقنن لعينة من السلوك".

يعرف كرونباخ Cronbach 1966 الاختبار تعريفاً عاماً ويعتبره "أي طريقة منظمة للمقارنة بين سلوك فردين أو أكثر".

إن معنى ذلك يشير إلى أن عينة من السلوك تعني أن الاختبار النفسي يقوم على ملاحظة وقياس عينة صغيرة جيدة الانتقاء من سلوك الفرد، فإذا أراد

السيكولوجي أن يختبر المحصول اللفظي للطفل، فإنه يفحص أداء الطفل في "عينة" ممثلة من الكلمات والألفاظ.

وقد لا يكون الاختبار بالضرورة عينة سلوك في كل الحالات، ففي ميدان الاختبارات التحصيلية هناك ما يسمى بالاختبارات المرجعية إلى المحك، وتقوم على أساساً على ملاحظة معظم أو جميع السلوك وليس عينة منه. حيث يوضع الاختبار بحيث يشمل ما تم تدريسه في الوحدة الدراسية من محتوى وأهداف وليس عينة منهما.

أما الجمعية السايكلوجية البريطانية 1966 فقد أصدرت نشرة تتضمن رأياً يتعلق بتعريف الاختبارات السايكلوجية يدعو إلى قصرة استخدام لفظ (اختبار) على أساليب التقدير التي تمدنا بتقديرات أو درجات نتيجة تطبيق طرق توصف بدقة في الكتيب المصاحب للاختبار ويوضح الكتيب بدقة الطريقة المستخدمة في تقنين الاختبار وبهذا فإن هذا التعريف يستبعد أي أسلوب قياسي غير مقنن ولا يعطي أي تقديرات نوعية وأحكام ذاتية.

ويشير لوفيل ولوسون 1976 بأن الاختبار المقنن (Standardized test) هو اختبار له تعليمات محددة لتطبيقه وتصحيحه ويتكون من عدد محدود من الوحدات أو البنود وطبق على عينات متماثلة للمجتمع الأصلي الذي قصد أن يكون الاختبار له بفرض تحديد معايير له وتتيح طريقة تطبيق الاختبار ومحتواه تطبيق نفس الاختبار على أفراد في أماكن وأوقات مختلفة وبهذا يمكن مقارنة درجة فرد ما في اختبار مقنن بدرجات أفراد آخرين أخذوا نفس الاختبار وعادة ما تكون المعايير والإنجازات النمطية لاختبار ما هي متوسطات الدرجات التي حصلت عليها مجموعات ممثلة في مستويات عمر مختلفة، ويمكن تحديد هذه المعايير لمجموعات غير مجموعات العمر مثل المجموعات المهنية، أو مجموعات الخبرة.

وتعرف الغريب الاختبار المقنن هو الاختبار الذي صيغت مفرداته وكتبت

تعليماته بطريقة تضمن ثباته إذا ما كرر كما تضمن صدقه في قياس السمة أو الظاهرة التي وضع لقياسها.

مفهوم التقويم:

أجاز مجمع اللغة العربية في القاهرة أن يقال "قيمت" الشيء تقييماً بمعنى حددت قيمته وقدره، وذلك للتفرقة أو إزالة اللبس بين هذا اللفظ وبين "قومته" بمعنى طورته وعدلته وجعلته قويماً أو مستقيماً، وهذا يعني أن التقويم يتضمن في ثناياه الحكم على الشيء، ويتجاوز حدود هذا الحكم إلى التحسين والتطوير. وفي هذا يقول الخليفة العادل: "من رأى منكم اعوجاجاً فليقومه".

في المجال التربوي: يمثل إصدار الحكم على مدى تحقيق الأهداف التربوية ما يطلق عليه التقييم التربوي، وهذا يمكن أن يتبعه إجراء عمليا يتعلق بتحسين وتطوير العملية التربوية، قد يكون في صورة برنامجا علاجيا مثلا، وهذا كله بالطبع يتجاوز معنى (التقييم التربوي) إلى معنى (التقويم التربوي).

وهذا يعني أن التقويم التربوي في معناه التربوي الواسع يقصد به: عملية منظمة لجمع وتحليل المعلومات، بغرض تحديد درجة تحقيق الأهداف التربوية، واتخاذ القرارات بشأنها لمعالجة جوانب الضعف وتوفير ظروف النمو السليم المتكامل من خلال إعادة تنظيم البيئة التربوية وإثرائها.

وعلى مستوى المدرسة يقصد بالتقويم: عملية منظمة لجمع وتحليل المعلومات حول البرامج المتعلقة بالمتعلم والمعلم والإدارة والأهداف ومحتوى المقررات والوسائل والنشاطات التعليمية والمرافق، وذلك لمعرفة مدى تحقيق الأهداف واتخاذ القرارات بشأن هذه البرامج.

أما على مستوى غرفة الصف فإن التقويم: عبارة عن عملية منظمة لجمع وتحليل المعلومات لتحديد مدى تحقيق الأهداف التعليمية من قبل المتعلمين، واتخاذ

قرارات بشأنها. ويشير هذا التعريف ضمناً إلى ضرورة صياغة الأهداف التعليمية: المعرفية، والنفسحركية، والوجدانية، كخطوة أولى في عملية التعليم لأنها عملية هادفة.

وفي المجال النفسي: يمثل إصدار الحكم على طفل معين بأنه متخلف عقلياً مثلاً يدخل في باب (التقييم النفسي) إن لم يتجاوز إصدار الحكم على المستوى العقلي، وهذا أمر نادر في الممارسة السيكولوجية، لان الشائع أن يتبع ذلك (إجراء عملياً) من نوع ما كإلحاق الطفل بمدرسة خاصة وتعليمه بطرق تتناسب مع مستواه العقلي وكل هذا من نوع "التقويم النفسي" أي أن التقويم النفسي هو القاعدة، والتقييم النفسي هو الاستثناء.

ما أنواع التقويم؟

للتقويم عدة تصنيفات نكتفي منها بتصنيفه تبعاً لتوقيت إجراءه في العملية التعليمية:

1- التقويم القبلي:

الأهداف التعليمية هرمية، بمعنى أن تحقيق بعضها يحتاج إلى تحقيق أهداف أخرى سابقة لها.، ولذلك فإن إجراء تقويم قبلي (تقويم تمهيدي) يساهم في:

- إعادة النظر في الأهداف التي لم يتمكن المتعلمين من متطلباتها السابقة.
- الكشف عن الأهداف التي يتقنها المتعلمين قبل البدء في التدريس، مما قد يترتب عليه ترك وحدة كاملة والانتقال إلى وحدة أخرى.
- البدء مع كل تلميذ من حيث مستواه.
- قد يؤدي إلى اتخاذ قرار بتقسيم المتعلمين إلى مجموعات أكثر تجانساً حسب درجة التمكن أو حسب أسلوب التعلم.

2- التقويم التكويني (البنائي):

هو التقويم الذي يجرى أثناء عملية التعليم والتعلم للتأكد من سلامة سير هذه العملية طبقاً للأهداف المرسومة لها، وتوجيهها.

ويساهم التقويم التكويني في تزويد المعلم والمتعلم بالتغذية الراجعة المتعلقة بالنجاح والفشل، فالمتعلم يشعر بنجاحه ويحدد أخطائه، والمعلم يعدل خططه على ضوء النتائج.

3- التقويم التشخيصي:

يستخدم هذا التقويم مع بعض المتعلمين الذين لا يحرزون تقدماً مرضياً في الاختبارات التي تطبق عليهم، ولا يبدون تحسناً بالرغم من إجراء تدريس علاجي لهم، وقد يدل ذلك على وجود صعوبات في التعلم نتيجة لأسباب قد تكون جسمية أو عقلية أو نفسية.

وقد يحتاج المعلم هنا إلى الاستعانة بالمرشد النفسي أو الأخصائي الاجتماعي في المدرسة، وقد يتطلب الموقف تطبيق بعض الاختبارات النفسية (كاختبارات القلق ومفهوم الذات)، أو العقلية (كاختبارات الذكاء).

4- التقويم الختامي (الإجمالي):

يتم التقويم هنا لنواتج التعلم في نهاية وحدة أو فصل أو سنة دراسية. وغالباً يتم رصد درجات المتعلمين باستخدام الاختبارات التحصيلية والتي يتم إعدادها من قبل المعلم، أو من قبل فريق من المعلمين.

ويتوقع أن يحقق هذا التقويم عدة أهداف رئيسة وثانوية، وقد يطغى الهدف الثانوي على الرئيسي أحياناً فقد يكون الهدف مقارنة تحصيل المتعلمين في عدة مدارس، أو الكشف عن مدى فعالية العملية التعليمية، باعتماد الدرجات كم} شر على مدى تحقيق الأهداف التعليمية.

مفهوم التقويم والقياس والعلاقة بينهما:

يصدر كل واحد منا يومياً أحكاماً كثيرة بشكل تلقائي أو طارئ أو منظم خلال مواقف حياتنا اليومية فنقول هذا جيد وهذا رديء وذاك أطول أو أقصر أو أحسن أو أفضل وهكذا.

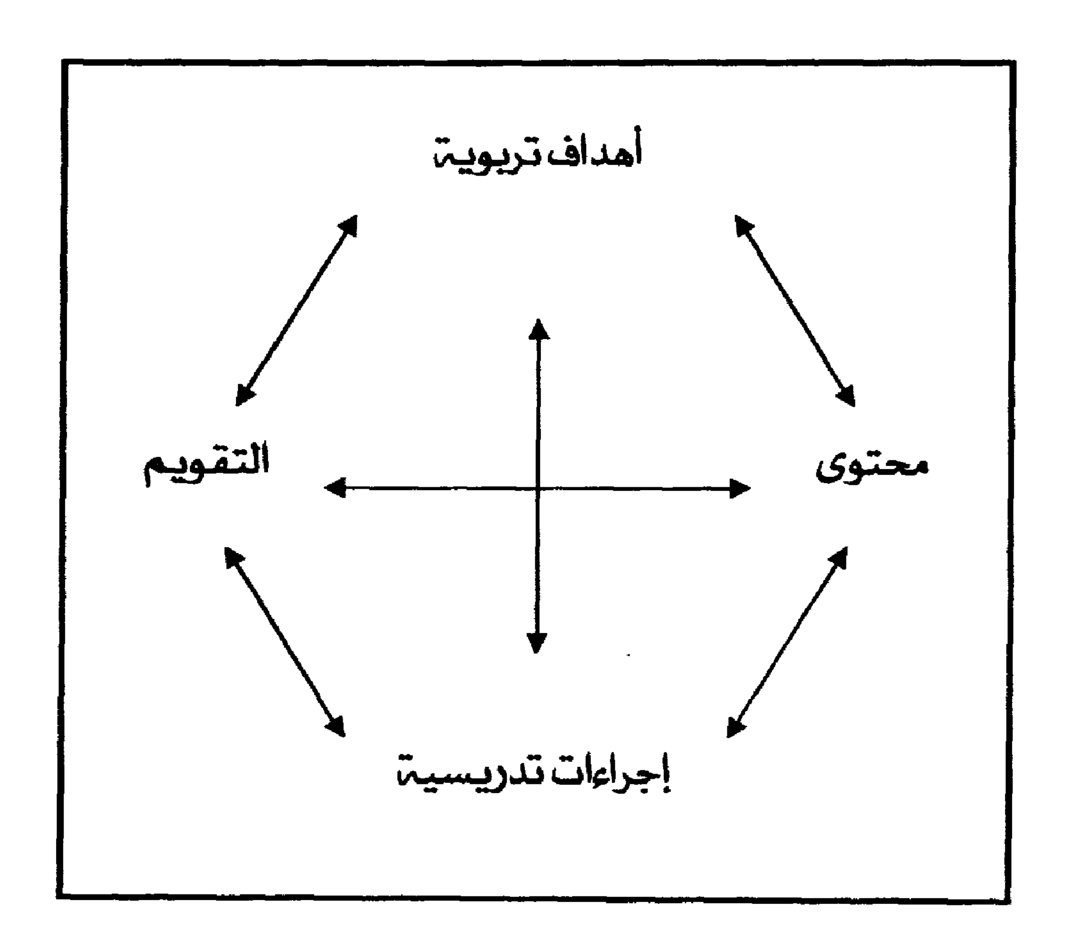
فالتقويم عملية طبيعية ودائمة وتشكل أساس تكيف الإنسان وتفاعله مع بيئته فهو باستمرار يقوم بتقويم عناصر حياته: يحدد ويقيس ويعدل ويفاضل ويزن ويقارن ويحكم ويختار في أفعاله وعن نفسه وعلى أفعال الآخرين وشخصياتهم وجميع هذه العمليات التقويمية قد تعتمد على الملاحظة أو على مقاييس تختلف من حيث دقتها وموضوعيتها. ويحدث كثيراً أن يخطئ الإنسان في أحكامه أو في تقويمه ولا يعي مصدر الخطأ مباشرة ويحدث أيضاً أن تكون تقديراته صحيحة دون أن يدرك العوامل التي ساعدت على صدق أحكامه، وكلما ازدادت معلومات الإنسان حول الموضوع المراد تقويمه، أزداد احتمال أن يكون حكمه أقرب إلى الصحة. أما حينما يخطأ الإنسان فإن الأحكام التي يصدرها أما أن تكون متسرعة أو غير دقيقة أو غير موضوعية. وهي في الحقيقة أحكام ناقصة أو غير صحيحة وتفتقر إلى الدقة والموضوعية.

والتقويم قد يكون ذاتياً وغير موضوعي والفرق الأساسي بينهما هو أن التقويم الموضوعي يخضع إلى الملاحظة الدقيقة المضبوطة ويعتمد على معلومات صادقة وثابتة ويقوم به عادة مختص يستعمل أدوات قياس دقيقة أما التقويم غير الموضوعي فيعتمد على الملاحظة العابرة أو التخمين أو الحدس السريع والتعميمات المطلقة والقياس غير الدقيق.

ومما تقدم يمكن تلخيص مفهوم التقويم بأنه عملية توفير معلومات موضوعية (صادقة وثابتة) لأجل إصدار حكم.

التقويم والعملية التعليمية:

يعد التقويم أحد أركان العملية التعليمية والشكل التالي يوضح مكانة وأهمية التقويم بين الأركان الأخرى للعملية التعليمية:



يتضح من المخطط أن لكل عملية تعليمية أهدافها، وفي ضوء هذه الأهداف يتم اختيار المناهج التعليمية التي تشكل محتوى هذه الأهداف، كما أن لهذه العملية التعليمية مجموعة من الإجراءات والأنشطة والأساليب التي يتم اختبارها لتحقيق الأهداف في ضوء المحتوى. وبعد انتهاء الإجراءات التربوية لابد من تقويم نواتج التعليم لدى المتعلمين والذي في ضوء نتائجه يتضح أولاً مدى تحقق الأهداف. كما تشكل نتائج التقويم إطارا مرجعياً لكل عمليات التحسين والتطوير التي يجب أن تصاحب العملية التربوية.

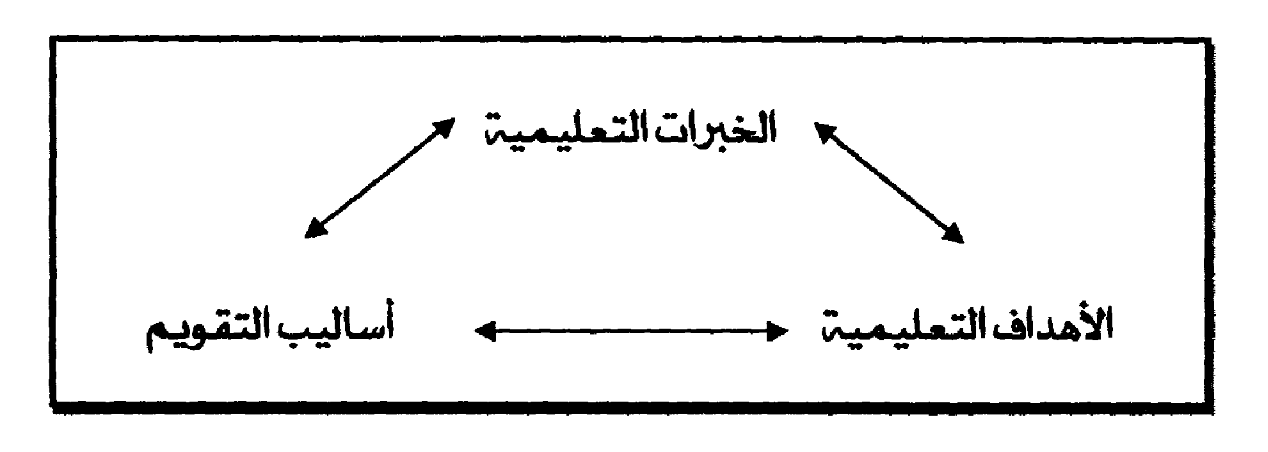
وتمثل عملية التقويم بالنسبة للمعلم من خلال المعلومات التي يجمعها عن تحصيل التلاميذ عاملاً مساعداً على التخطيط السليم للخبرات التعليمية المناسبة، وتحديد مدى تحقق تلامذته للأهداف التعليمية، وتمكنه من تقويم أسلوبه وطريقته في التدريس. وبالتالي فإن القرارات التي يتخذها المعلم في سبيل تحسين هذه العملية تكون مبنية على أسس وأساليب صحيحة. ولتحقيق ذلك يجب على المعلم أن يقوم بالخطوات الآتية:

أولاً: تحديد الخبرات التعليمية التي يحويها المقرر الدراسي تحديداً دقيقاً.

ثانياً: وضع الخبرات التعليمية في صورة أهداف سلوكية محددة وواضحة.

ثالثاً: تحديد أساليب وإجراءات التقويم المناسبة لمعرفة الدرجة التي تحققت بها هذه الأهداف واتخاذ القرارات المناسبة لتحسينها وتطويرها.

وهذه الخطوات الثلاثة يمكن تمثيلها في الشكل التالي:



نلاحظ من التخطيط السابق أن الخطوات الثلاثة مرتبطة ببعضها البعض تؤثر كل منها بالأخرى وتشير إلى الطبيعة التفاعلية للعملية كلها، فإذا كانت إجراءات التقويم ضعيفة، فإن المعلومات التي تبنى عليها أحكامنا لن تكون كافية.

إن عملية التقويم كما ذكرنا سابقاً تتصف بالتنظيم والشمول، وهي تشمل كل العناصر المرتبطة بالعملية التعليمية ويسهم في جمع معلوماتها وبياناتها كل الفعاليات المشاركة فيها. ولذلك تتسع مجالات التقويم لتشمل كل جوانب هذه العملية التربوية.

استخدام الاختبارات:

استخدمت الاختبارات والمقاييس النفسية والتربوية للأغراض التالية:

1- تقويم التعلم:

تهدف الاختبارات المدرسية إلى معرفة المدى الذي وصل إليه التلميذ في تعلم المادة المعينة. إن مثل هذه الاختبارات تحاول تقويم التعلم ويمكن أن يعيد المعلم النظر في طريقة تعليمه للمادة الدراسية في ضوء النتائج التي يحصل عليها فيما إذا حصل الطلاب على درجات منخفضة مفترضين أن المعلم قادر على تعيين مستوى الصعوبة (صعوبة المادة) بحيث يحصل الطالب المتوسط على درجة متوسطة وهكذا. إن نسب النجاح قد تكون مؤشراً من بين مؤشرات عديدة ولكنها غير كافية لوحدها. وثمة اختبارات يستطيع المعلم أن يختبر نفسه بها ذاتياً، أو أن يقوم شخص آخر كالمدير أو آخرين بإجراء الملاحظة وفق اعتبارات منهجية خاصة.

2- قياس أساليب أداء المفحوصين وإمكاناتهم:

أ- تقيس اختبارات الذكاء مثلاً القدرة العقلية العامة وتقيس اختبارات متخصصة معينة مدى محدوداً من القدرات مثل تلك القدرات اللازمة للفهم الميكانيكي أو الحكم على العلاقات المكانية.

ب- وتقيس اختبارات المعلومات والتحصيل مستوى الكفاية الحالية في موضوع أو مهارة حصلها الفرد كنتيجة لتعليم خاص وهي تتضمن اختبارات الكفاية في الكتابة على الآلة الكاتبة والهجاء والقراءة والحساب وكذلك بطارية التحصيل التربوي التي تقيس الأداء في مجالات عديدة.

ج- وتتنبأ اختبارات الاستعدادات بقدرة المفحوص على أن يحسن أداءه بالتدريب الإضافي في ميدان أكاديمي أو مهني خاص. ورغم أن هذه الاختبارات تقوم على الأداء الراهن. إلا أنها تقيس عادة بعض المجالات التي لم يتلق المفحوص

فيها تدريباً خاصاً. وقد تقيس عوامل مثل المهارة الميكانيكية أو التناسق الحركي أو القدرة الموسيقية أو الفنية أو تقيس الاستعداد للطب أو الهندسة أو اللغات أو الجبر.

3- قياس تفضيلات الأفراد وسلوكهم:

أ- بعض هذه الاختبارات تقيس ميول المفحوصين نحو مهن أو أنشطة معينة

ب- ويتعرف بعضها الأخر على طبيعة وأبعاد الاتجاهات والمعتقدات التي يتمسك بها الأفراد أو الجماعات إزاء القضايا الجدلية في المجتمع وأنشطته ومؤسساته وقطاعا ته.

ج- وتقيس نوع من الأدوات العوامل الانفعالية والاجتماعية، مثل تكيف الفرد مع نفسه ومع الآخرين.

د- وتقيس الاختبارات جوانب من سلوك المفحوص مثل سلوكه الأخلاقي أو صفات التعاون أو الصداقة أو القيادة عنده.

هـ- قياس وتحديد الأوضاع البيئية والمادية للأفراد والمؤسسات. إن بعض هذه المقاييس يقيس نواحي معينة في المنزل مثل الوضع الاجتماعي الاقتصادي للأسرة أو علاقات الأب، الأم، أو الطفل أو الوالد، الطفل. وتقوم وسائل مسحية متنوعة عوامل معينة في مدرسة أو مؤسسة أو مجتمع محلى مثل عدد الخدمات ونوعها وخدمات القيادة والبرامج. والأنشطة والتماسك. وتقيس اختبارات اللياقة الطبية والجسمية الحالة الصحية للأفراد.

4- تصنيف التلاميذ:

ثمة حاجة لتصنيف الأطفال قبل دخولهم المدرسة، حيث يحدث أن يختار المسؤولون التربويون تصنيف بعض الأطفال الذين يعانون من ضعف القدرات العقلية تبدو في أعراضها كالتخلف العقلي، وقد يحدث العكس، حيث يقبل تلاميذ

في المدرسة بينما هم يعانون من تخلف عقلي ومثل هذا الخلط يمكن أزالته والتأكد من استعداد التلميذ الدراسي باستخدام الاختبارات العقلية التي تقيس الذكاء العام أو الاستعداد - الدراسي.

5- قياس تحصيل التلاميذ:

قد تظهر الاختبارات التي يعدها المعلمون ضعفاً في ثباتها وأحياناً في صدقها ولكنها بتنوعها وتعدد مناسباتها يمكن أن تقيس تحصيل الطلبة. فالامتحانات اليومية والشهرية والفصلية والسنوية يمكن أن تعطي صورة عن تحصيل التلاميذ لكنها لا تخلو من أخطاء في دقتها، بينما تستطيع الاختبارات التحصيلية المقننة والمعدة من قبل جهة متخصصة قياس تحصيل التلاميذ بدرجة عالية من الموضوعية. ومثل هذه الاختبارات تكون لها استخدامات مهمة في تقدير المستوى الحقيقي لمجموعة من الطلبة في مادة دراسية معينة وقد تتم الاستفادة منها في تقرير مستوى التدريس والأساليب المناسبة للتعليم

6- تحديد صعوبات التلاميذ في التعلم:

تكشف الاختبارات دائماً عن عوامل الضعف في تحصيل التلاميذ أو الصعوبات التي يعانون منها من خلال تحليل الإجابات الخاطئة. ومنها يستطيع المعلم التصدي لأخطاء التلاميذ وصعوباتهم التعليمية.

7- تقويم الأهداف السلوكية:

تستطيع الاختبارات أن تكشف لنا عن التغير الحاصل في اتجاهات التلاميذ وقدراتهم في المجالات الانفعالية أو العقلية أو في المهارات العملية والتي وضعت كأهداف سلوكية مرغوبة، ومثل هذه التغيرات قد تصعب ملاحظاتها أو تقديرها إلا باستعمال الاختبارات النفسية المعدة خصيصاً لمثل هذا المقياس. كما أن هذه الاختبارات تساعد المعلمين على معرفة الفروق الفردية بين التلاميذ

8- تيسير عمليات التوجيه التربوي والإرشاد:

حيث تكون الاختبارات أدوات فعالة وأساسية في تعريف الطلبة بقدراتهم وتوجيههم بما يتناسب بتلك القدرات. كما يمكن استعمالها في عمليات الإرشاد النفسي والتوجيه المهني في قياس بعض الصفات الشخصية وتشخيص المشكلات التي يعاني منها الطلبة، أو تشخيص قدراتهم واستدلالهم وميولهم وتوجيههم وفقاً لها.

9- تقويم الكفاءة الإدارية للمدرسة:

حيث يمكن للاختبارات أن تؤدي دوراً مهماً في تقويم الأنشطة الإدارية وعناصرها وخاصة توضيح القيادة التربوية داخل المدرسة ضمن شروط وتعليمات خاصة من ذلك الاختبارات التي يمكن أن تقيس السمات الواجب توفرها في اختيار المدير الناجح في المدرسة.

10- تقويم المنهج والأنشطة المدرسية المختلفة:

يمكن للاختبارات أن تقيس درجة كفاءة الكتب والمفردات المقررة وتحديد صعوبات التلاميذ فيها أو تقويم الأنشطة والفعاليات المدرسية مثل الوسائل التعليمية والأبنية المدرسية.. الخ.

11- مجال البحث العلمي:

وتستخدم الاختبارات كأدوات في إجراء البحوث والدراسات النظرية كقياس الذاكرة ومعرفة أثر التكرار المنظم في الذاكرة أو دراسة أثر عوامل البيئة على تحصيل التلاميذ، أو العلاقة بين الإبداع وأنماط التربية الأسرية والى غيرها من الاستخدامات.

أنواع الاختبارات:

هناك أنواع عديدة جداً من الاختبارات النفسية والتربوية إذ يتجاوز عددها اليوم أكثر من عشرة آلاف اختبار مقنن تستعمل في أغلب الدول المتقدمة وتخص

هذه الاختبارات أغلب عناصر السلوك البشري والجوانب التربوية المرتبطة بها، لذا فإن تصنيفها ليس هيناً وثمة طرق مختلفة لتصنيف هذه الاختبارات، فقد صنفت حسب:

1- التصنيف حسب درجة تحديد المثير والاستجابة:

أ- اختبارات إسقاطية: وهي الاختبارات التي لا يكون فيها المثير محدد، ولا الإجابة محددة، كالاختبارات النفسية التشخيصية مثل اختبار رورشاخ (أو اختبار بقع الحبر).

ب- اختبارات محددة البناء: هي الاختبارات التي يكون فيها المثير واضحاً، أو يكون المطلوب في السؤال محدداً، كما أن هناك مفتاح إجابة محددة، كاختبارات التحصيل والاستعداد.

2- التصنيف حسب طبيعة الأداء:

أ- أقصى الأداء: وهي الأدوات التي يتم فيها أثارة دافعية المتعلم لتقديم أفضل ما عنده من إجابة والحصول على أعلى درجة، مثل الاختبارات التحصيلية بأنواعها، ونهيئ الفرصة له للدراسة والاستعداد، وتحقيق أعلى مستوى للتحصيل.

ب- الأداء العادي أو الطبيعي: هي الأدوات التي تعكس سلوك المتعلم في الظروف العادية أو الطبيعية دون محاولة خارجية لتوجيه هذا السلوك حتى يكون التقدير الذي يحصل عليه منسجماً مع السلوك الواقعي. فإذا كانت الأداة تقيس قلق المتعلم في الامتحان مثلاً فإن القرارات التي يتم اتخاذها في ضوء النتائج تعتمد على صدق الإجابات ولذلك تعتبر مقاييس الاتجاهات والميول من نوع الأداء العادي.

3- التصنيف حسب عدد الأفراد الذين يطبق عليهم الاختبار:

أ- اختبارات فردية: وهي التي لا تطبق إلا على تلميذ واحد في المرة الواحدة مثل الاختبارات الشفوية أو اختبارات الذكاء مثل اختبار ستانفورد بينيه.

ب- اختبارات جماعية: وهي التي تطبق على عدد كبير من المتعلمين بنفس الوقت.

4- التصنيف حسب سرعة الإجابة:

أ- اختبارات السرعة: وهي الاختبارات التي تكون فيها سرعة الإجابة هي العامل الحاسم في تحديد أداء المتعلم، حيث يصعب على معظم التلاميذ إنهاء الإجابة عن جميع الفقرات ضمن الزمن المحدد.

ب- اختبارات القوة: وهي الاختبارات التي يعطى فيها زمن شبه مفتوح للإجابة بحيث يكون كل تلميذ قادراً على محاولة الإجابة عن كل سؤال، إلا أن صعوبة الأسئلة وقدرة المتعلم على إجابتها هي التي تحدد أداءه، بمعنى أن المتعلم لا يحصل على الدرجة النهائية بسبب صعوبة الأسئلة، وليس بسبب كثرتها.

5- التصنيف حسب طريقة تفسير النتائج:

صنفت أدوات القياس حسب طريقة تفسير النتائج إلى فئتين هما:

أ- الاختبارات معيارية المرجع NRT: ذكر بابام وهيوسك (Pohpam & Husek 1969) بأن هذه الاختبارات تعتمد عند تفسيرها للدرجات المستحصلة على خصائص المجموعة التي طبقت عليها حيث يجب أن تكون هذه الخصائص مماثلة مع خصائص الطلاب من حيث السن والنوع والمنطقة الجغرافية والمرحلة الدراسية، وتهتم هذه الاختبارات بالكشف عن الفروق الفردية بين الطلاب وذلك من خلال مقارنة أداء الطالب بأداء أقرانه كجماعة معيارية بمعنى أن هذه الاختبارات تركز على وضع الطالب النسبي في علاقته بالآخرين.

ب- الاختبارات محكية المرجع CRT: أشار بابام وهيوسك (Pohpam & Husek

1969) بأن هذه الاختبارات لا تعتمد عند تفسيرها لدرجة الطالب على خصائص مجموعته التي طبق عليها الاختبار، ولكنها تعتمد على مقارنة أداء الطالب بمحك أو مستويات أداء متوقعة ومحددة مسبقاً ويتم تحديد هذه المستويات في ضوء الأهداف السلوكية (الإجرائية) المراد قياسها وتقيس هذه الاختبارات مدى تحقيق الطالب لهذه الأهداف السلوكية وبذلك يمكن التحقق من كفاءة كل طالب من اكتسابه للمهارات والمعارف المرجوة ومن ثم تشخيص نواحي الضعف والقوة.

6- التصنيف حسب شكل (نوع) الفقرة:

صنفت أدوات القياس والتقويم حسب شكل (نوع الفقرة) إلى عدة أشكال من أهمها: المطابقة (المزاوجة)، والصواب والخطأ، والاختيار من متعدد، والتكميل، والأسئلة المقالية (الإنشائية).

7- التصنيف حسب الجهة التي تعد الاختبار:

أ- اختبارات من إعداد المعلم: وهي اختبارات غير رسمية، حيث يعدها المعلم نفسه.

ب- اختبارات مقننة (رسمية) أو منشورة: حيث يعدها فريق من المختصين في مراكز القياس والاختبارات.

8- التصنيف حسب الكيفية التي يظهر فيها الأداء (أسلوب تقديم الإجابة)

أ- اختبارات لفظية: هي الاختبارات التي تكون فيها الإجابة تحريرية (ورقة وقلم) أو شفوية.

ب- اختبارات الأداء المبرهن عملياً: هي التي تتم في المختبرات، مثل استعمال جهاز، أو إجراء تجربة عملية، أو في الجانب العملي للتعليم المهني.

إن التصانيف أعلاه من الأدوات يمكن أن تتبع لأكثر من تصنيف، فاختبار التحصيل يمكن أن يكون اختبار قوة، ويفسر تفسيراً محكي المرجع، ومن نوع أقصى الأداء.

العوامل التي تؤثر في القياس:

يتأثر القياس بعوامل مختلفة نذكر منها:

1- الشيء المراد قياسه أو السمة المراد قياسها: يؤثر في نوع القياس، والأداة المستخدمة في القياس، فهناك أشياء تقاس بطريقة مباشرة كأطوال المتعلمين عند الكشف الطبي عليهم، ولكن الاستعدادات العقلية والسمات المزاجية والشخصية تقاس بطريقة غير مباشرة، ولذلك اختلفت مقاييسها عن مقاييس الطول في طبيعتها ودرجة دقتها، وليس من شك أن القياس المباشر أسهل وأدق من القياس غير المباشر.

2- أهداف القياس: تؤثر في الطريقة التي نستخدمها في القياس، في حين يكون الهدف من القياس عمل تقويم سريع لتحصيل المتعلمين في خبرة معينة مثلاً، اختيرت الطريقة التي تناسب هذا الهدف. وحين يكون هدف القياس عمل تقويم شامل ودقيق لظاهرة معينة رسمت خطة دقيقة للقياس واختير المقياس بدقة، وكذلك القائمون بالقياس ممن تدربوا على العملية تدريباً دقيقاً.

3- طبيعة الظاهرة أو السمة المقاسة: فبعض السمات يمكن التحكم فيها وقياسها بدقة مثل الذكاء والاستعدادات العقلية، ولكن يصعب التحكم في سمات أخرى كالسمات المزاجية وسمات الشخصية، كما يصعب تحديدها وتصميم المواقف التي تمثلها بدقة بسبب تعقدها، وتأثرها بعوامل ذاتية عديدة يصعب استبعادها.

مبادئ عامة ينبغي مراعاتها في التقويم:

1- ضرورة تحديد الغرض من التقويم: إذا لم يكن الغرض من التقويم واضحاً، فإنه من الصعب الحكم على جدوى عملية التقويم، كما يصعب التأكد من صحة الخطوات اللاحقة لهذه العملية، كاختيار أسلوب التقويم المناسب لجمع المعلومات. فهدف ترتيب المتعلمين طبقاً لتفوقهم، يختلف عن هدف التنبؤ بمدى نجاحهم في تخصص معين مستقبلا "، وهذا يختلف بدوره عن هدف تشخيص جوانب الضعف في مهارات القراءة والكتابة والحساب.

2- الاهتمام باختيار وأعداد أدوات التقويم المناسبة: وينبغي أن تتسم بالصدق والثبات والموضوعية.

3- ضرورة وعي المعلم المقوم بمصادر الأخطاء المحتملة في عملية التقويم ومنها:

أ- خطأ العينة: وهو الخطأ الذي ينتج من عدم تمثيل عينة الأسئلة في الاختبار، فقد تكون الأسئلة غير كافية من حيث العدد، أو من حيث تغطيتها للموضوعات الدراسية، وقد تكون الأسئلة متحيزة لوحدة معينة على حساب الوحدات الأخرى، وربما كانت متحيزة لمستوى عقلي معين (كالتذكر في مستويات بلوم) على حساب المستويات الأخرى.

ب- أخطاء التخمين والتورية: يقصد " بالتخمين " هنا تخمين المتعلم للإجابة، ويرتبط هذا الخطأ بالأسئلة الموضوعية. أما خطأ التورية فينتج عند محاولة المتعلم كتابة أي شيء عندما لا يعرف الإجابة فعلاً في حالة الأسئلة المقالية.

ج- أخطاء التحيز الشخصي أو أثر الهالة: وقد ينتج هذا الخطأ من تكوين المعلم مثلاً صورة معينة عن تلميذ في موقف معين بشكل يؤثر على تقديره له في مواقف أخرى، وقد يكون الأثر سلبياً أو إيجابياً.

د- أخطاء التزييف والرغبة الاجتماعية: وتنتج هذه الأخطاء من محاولة المفحوص أن يزيف أجابته بحيث يعطي صورة معينة عن نفسه مغايرة للحقيقة، فقد يرتاح المتعلم عندما يعكس نفسه بصورة شخصية متواضعة وهو عكس ذلك، أو يظهر اتجاهاً إيجابياً نحو مادة معينة في مقياس الاتجاه نحو التخصصات المختلفة لأغراض التوجيه والإرشاد.

هـ- أخطاء البنية الشخصية: وقد ينتج هذا الخطأ من البنية الشخصية للمعلم، أو البنية الشخصية للتلميذ. فقد يتصف معلم بالليونة، وآخر يتصف بالقسوة، وثالث يميل نحو الوسطية ويظهر ذلك جلياً في تقديره لإجابات المتعلمين الإنشائية وفي التقارير.

أما بالنسبة للتلاميذ فعند تطبيق مقاييس التقدير عليهم (مثل: موافق بشدة، موافق، حيادي، أعارض، أعارض بشدة) بعضهم يميل إلى الحياد، وبعضهم إلى الوسطية، والبعض يميل إلى التطرف.

4- الوعي بخصائص عملية التقويم: وأهم هذه الخصائص الشمولية لجميع أوجه النمو والخبرات، والاستمرارية.

5- التأكد من أهمية البرنامج المقوم والالتزام بأخلاقيات عملية التقويم. ويمكن توضيح المقصود من خلال تقويم التحصيل الأكاديمي، فالمعلم يجب أن يكون حريصاً على أن يضمن اختباراته الأسئلة التي تقيس الأهداف الأساسية بدلا من التركيز على الأهداف الهامشية.

وقد يوجه التقويم سلفاً للخروج بنتيجة إيجابية سواء من خلال إهمال جوانب أساسية في موضوع التقويم، أو التقليل من وزنها، أو من خلال الاختيار السيئ لفريق التقويم نفسه

وقد لا يلتزم المقوم بأخلاقيات عملية القياس والتقويم، كأن لا يحافظ على سرية المعلومات إذا تطلب الأمر ذلك، أو يستخدمها في التشهير والتجريح والتوبيخ، وهذه نقطة هامة بالنسبة للمعلم وللمرشد النفسي والأخصائي الاجتماعي.

6- ضرورة احتواء تقرير التقويم على معلومات كاملة وافية بشكل يمكن صانعي القرار من اتخاذ القرار المناسب أو الاختيار من بين بدائل القرارات.

الفصل الثاني

الأهداف التعليمية والسلوكية

يحتوي الفصل على:

- تعريف الهدف السلوكي
- أهمية تحديد الأهداف
- هل الهدف ضرورة ملحة في العملية التربوية؟
- إيجابيات الأهداف في العملية التربوية
- تصنيف بلوم للأهداف التعليمية
- تصنيف زايس
- تصنيف انيتاهارو
- صياغة الأهداف السلوكية
- اعتبارات ينبغي مراعاتها عند الآخذ بالأهداف السلوكية
- دور الأهداف السلوكية في العملية التعليمية

تعريف الهدف السلوكي

يعرف "كمب" الهدف السلوكي بأنه عبارة دقيقة تجيب عن السؤال التالي: ما الذي يجب على الطالب أن يكون قادرا على عمله ليدل على إنه قد تعلم ما تريده أن يتعلم.

ويعرف المركز العربي للبحوث التربوية لدول الخليج العربي الهدف السلوكي بأنه التغير المرغوب المتوقع حدوثه في سلوك المتعلم والذي يمكن تقويمه بعد مرور المتعلم بخبرة تعليمية معينة.

ويعرف الصالح الهدف السلوكي بأنه وصف دقيق وواضح ومحدد لناتج التعلم المرغوب تحقيقه من المتعلم على هيئة سلوك قابل للملاحظة والقياس.

أما مصطفى السيد فيعرفه بأنه(عبارة تصف التغير المرغوب فيه في مستوى من مستويات خبرة أو سلوك المتعلم معرفياً، أو مهارياً، أو وجدانياً عندما يكمل خبرة تربوية معينة بنجاح، بحيث يكون هذا التغير قابلاً للملاحظة والتقويم.)

أهمية تحديد الأهداف:

الأهداف دائما نقطة البداية لأي عمل سواء كان هذا العمل في إطار النظام التربوي أو أي نظام آخر، فهي تعد بمثابة القائد والموجه لكافة الأعمال. ويمكن إبراز الدور الهام للأهداف التربوية على النحو التالي:

1- تعنى الأهداف التربوية في مجتمع ما بصياغة عقائده وقيمه وتراثه وأماله واحتياجاته ومشكلاته.

2- تعين الغايات لمخططي المناهج على اختيار المحتوى التعليمي للمراحل الدراسية المختلفة وصياغة أهدافها التربوية الهامة.

3- تؤدي الأهداف التربوية دورا بارزا في تطوير السياسة التعليمية وتوجيه العمل التربوي لأي مجتمع.

4- يساعد تحديد الأهداف التربوية في التنفيذ الجيد للمنهج من حيث تنظيم طرق التدريس وأساليبها وتنظيم وتصميم وسائل وأساليب مختلفة للتقويم.

5- تساعد الأهداف التربوية على وضوح الرؤيا، فأي عمل ناجح لا بد من أن يكون موجهاً نحو تحقيق أهداف محددة ومقبولة، وإلا أصبح العمل نوعاً من المحاولة والخطأ التي تعتمد على العشوائية والارتجال وفي هذا ضياع للوقت والجهد والمال، مما لا تستطيع طاقتنا أن تقبله أو تتحمل عواقبه. هذه العشوائية في حقيقة الأمر هي ما نود أن نتجنبه في تدريسنا.

6- تحديد الأهداف ضروري لاختيار الخبرات المناسبة:

- ماذا نختار من المعرفة والخبرات الإنسانية المتفجرة المتراكمة (معلومات، مهارات، طرق تفكير، اتجاهات وعادات وميول وقيم) لنكسبه لأطفالنا؟
- أي الخبرات لازمة للتلميذ لكي تساعده على التكيف الإيجابي مع مجتمعه في حاضره ومستقبله؟ وأي الخبرات الماضية نتيحها لهم؟

7- تحديد الأهداف ضروري لاختيار أوجه النشاط التعليمي المناسبة: إذا كان من أهداف المعلم تزويد التلاميذ بمهارات أو طرق تفكير أو اتجاهات معينة، فإن ذلك سيساعد المعلم على اختيار أنواع النشاط التعليمي التي يمكن أن تحقق هذه الأهداف. ويحدث كثيراً أن نخطىء في تدريسنا في اختيار النشاط المناسب ربما لا لسبب إلا لأننا لا نعرف على وجه التحديد ما هدية الأهداف التي نطمح في الوصول إليها.

8- تحديد الأهداف ضروري للتقويم السليم: التقويم "عملية تشخيصية وقائية علاجية تستهدف تحسين عملية التعليم والتعلم"، وإذا كانت هذه هي وظائف عملية التقويم، فلا بد أن يكون هناك أساس نبني عليه أحكامنا، ولا يمكن أن نبني تقويم تدريسنا سوى على الأهداف التي نتفق عليها ونرتضيها. فالأهداف المحددة هي التي ستحدد "ماذاً نقوّم، وفي الوقت نفسه فإنها ترشدنا إلى اختيار أنسب الوسائل والأدوات التي تصلح لتقويم ما نهدف إليه.

هل الهدف ضرورة ملحة في العملية التربوية؟

إن الجواب على هذا السؤال يطرح جملة من القضايا الأساسية المرتبطة بأهمية الأهداف في العملية التربوية وهي قضايا يمكن إيجازها بالعناصر الآتية:

1- إن مفهوم التربية في جوهره يفيد تحقيق هدف ما.

2- إن ممارساتنا في الحياة اليومية في حد ذاتها مجموعة أهداف نسعى لتحقيقها.

3- إن الأهداف التربوية معيار أساسي لاتخاذ قرارات تعليمية عقلانية وعملية خاضعة للفحص والتجريب.

إيجابيات الأهداف في العملية التربوية:

إن هناك مجموعة من الإيجابيات يحققها التعليم بواسطة الأهداف ويمكن إيجاز تلك الإيجابيات بما يلي:

1- إن تحديد الأهداف بدقة يتيح للمعلم إمكانية اختيار عناصر العملية التعليمية من محتوى وطرائق ووسائل وأدوات تقويم.

2- إن تحديد الأهداف يسمح بفردانية التعليم.

3- إن تحديد الأهداف يساعد على إجراء تقويم لإنجازات التلاميذ.

4- إن المتعلم عندما يكون على علم بالأهداف المراد تحقيقها منه فإنه لا يهدر وقته وجهده بأعمال غير مطلوبة منه.

5- عندما تكون الأهداف محددة فإنه من السهل قياس قيمة التعليم.

6- إن وضوح الأهداف يضمن احترام توجهات السياسة التعليمية.

7- إن وضوح الأهداف يتيح إمكانية فتح قنوات تواصل واضحة بين المسؤولين على التربية والتعليم.

8- إن تحديد الأهداف يتيح للمتعلمين إمكانية المساهمة في المقررات على اعتبار أنهم يصبحون قادرين على تمييز التعليمات الرسمية وتقييمها.

9- إن وضوح الأهداف يتيح إمكانية التحكم في عمل التلميذ وتقييمه.

10- إن وضوح الأهداف يتيح إمكانية توضيح القرارات الرسمية لضبط الغايات المرسومة.

عمومية الأهداف وخصوصيتها:

تشكل الأهداف التربوية الغايات الأساسية التي نرغب من تلاميذنا بلوغها عند مرورهم بالخبرات التعليمية التعلمية في دروس المناهج المختلفة. وغني عن البيان أن الأهداف التربوية أما أن تكون عامة أو خاصة. فمن حيث هي أهداف عامة فهي تلك الأهداف التي يمكن أن يبلغها التلميذ في فترة زمنية طويلة ولا يمكن تحقيقها في حصة دراسية أو أسبوع دراسي مثل الأهداف العامة لتدريس العلوم في المرحلة الابتدائية العليا. فهذه الأهداف يمكن أن يبلغها التلميذ في نهاية منهج أو مناهج العلوم للمرحلة الابتدائية

أما الأهداف الخاصة فهي الأهداف المتعلقة بوحدة دراسية أو موضوع تعليمي، وهذا النوع من الأهداف يُطلق عليها أحياناً الأهداف السلوكية باعتبارها السلوكيات التي ترمي المدرسة إلى إكسابها لطلابها، أو الأهداف الأدائية باعتبارها ذلك الجزء من سلوك المتعلم الذي يمكن ملاحظته وتقويمه.

تصنيف بلوم للأهداف التعليمية

اتفق خبراء المناهج وطرق التدريس والقياس والتقويم على تصنيف الأهداف التعليمية في اجتماع لهم عام 1956 في جامعة شيكاغو إلي ثلاثة مجالات هي المعرفي Cognitive، والانفعالي Affective، والنفسحركي Psychomotor.

تأتي أهمية هذا التصنيف من صعوبة التعامل مع شخصية المتعلم المعقدة بصورة إجمالية، مع أننا نعرف أن الشخصية كل متكامل وفريدة في خصائصها، بمعنى أن الهدف من التصنيف التبسيط والتسهيل للمجالات لا الفصل بينها.

أولا: المجال المعرفي The Cognitive Domain

هو المجال الذي يتعلق بتذكر المعرفة، كما يمتد لتنمية القدرات والمهارات العقلية. ويصنف بلوم Bloom وزملائه هذا المجال إلى ست مستويات تتدرج من التذكر (المعرفة) إلى الفهم، ثم التطبيق، ثم التحليل، ثم التركيب، ثم التقويم.

1ـ مستوى المعرفة ـ التذكر Knowledge - Recall

التركيز هنا على تذكر حقائق فرع من فروع المعرفة، ومصطلحاته ومبادئه وتعميماته وقوانينه ونظرياته، ويكون هذا التذكر بنفس الصورة أو الصيغة التي عرض بها أثناء العملية التعليمية. ويتضمن هذا المستوى الجوانب المعرفية التالية:

- معرفة الحقائق المحددة: مثل معرفة أحداث محددة، تواريخ معينة، أشخاص، خصائص.
- معرفة المصطلحات الفنية. مثل معرفة مدلولات الرموز اللفظية وغير اللفظية.
- معرفة الاصطلاحات: مثل معرفة الاصطلاحات المتعارف عليها للتعامل مع الظواهر أو المعارف.
- معرفة الاتجاهات والتسلسلات: مثل معرفة الاتجاهات الإسلامية في السنوات الأخيرة بالغرب.
- معرفة التصنيفات والفئات.
- معرفة المعايير.
- معرفة المنهجية وطرائق البحث.
- معرفة العموميات والمجردات: مثل معرفة المبادئ والتعميمات ومعرفة النظريات والتراكيب المجردة.

أمثلة:

- أن يذكر التلميذ أسماء ثلاثة من العلماء العرب (أو المسلمين).

- أن يعدد التلميذ أسماء حقول البترول في العراق بنسبة صواب لا تقل عن 90 ٪.
- أن يكتب المتعلم أسماء الأنهار في العراق في زمن لا يزيد عن دقيقتين.
- أن يتلو المتعلم سورة الصافات دون أي أخطاء.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يذكر، يسمي، يتلو، يحدد، يتعرف على، يصف، يعدد.

2- مستوى الفهم- الاستيعاب Comprehension:

وهو القدرة على تفسير أو إعادة صياغة المعلومات التي حصلها الطالب في مستوى المعرفة بلغته الخاصة. والفهم في هذا المستوى يشمل الترجمة والتفسير والاستنتاج. ويستطيع المتعلم في هذا المستوى أن يعبر عما درسه من أفكار تعبيراً يختلف عما أعطي له، أو عُرض عليه أثناء الدراسة، ويشمل هذا أن يصيغ الفكرة بلغته، أو أسلوبه الخاص بشرط توفر الدقة والأمانة، أو أن يقوم بشرح وتلخيص وإعادة تنظيم الفكرة.

أمثلة:

- أن يُحوّل المتعلم العلاقة بين الحجم والكتلة والكثافة إلى علاقة رياضية.
- أن يُلخّص المتعلم خواص الهيدروجين التي قرأها في الكتاب المدرسي بلغته الخاصة.
- أن يعيد الطالب صياغة المقصود بيوم القيامة بأسلوبه الخاص.
- أن يصف المتعلم بلغته الخاصة العلاقة بين الحيوان والنبات بعد قراءة المقالة الواردة في المجلة العلمية.
- أن يعلل المتعلم أسباب انتصار المسلمين في معركة بدر الكبرى.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يُعبّر بلغته الخاصة عن، يوضّح، يفسّر، يناقش، يصيغ بأسلوبه، يعيد ترتيب، يستنبط، يستنتج، يلخّص.

3- مستوى التطبيق Application:

وهو القدرة على استخدام أو تطبيق المعلومات والنظريات والمبادئ والقوانين في موقف جديد.

ويستطيع المتعلم في هذا المستوى أن يستخدم ما درسه من معلومات في مواقف جديدة تختلف عن تلك التي تم فيها عرض المعلومات أثناء دراستها، ويشمل ذلك استخدام المجردات في مواقف واقعية.

الأسئلة في هذا المستوى قابلة للتضليل، بمعنى أنها قد تظهر بمستوى التطبيق ولكنها لا تتعدى حقيقة مستوى المعرفة، وقد يعود السبب إلى إعطاء المعلمين أسئلة مماثلة تماماً للأسئلة الموجودة في الكتاب المقرر. كما يحدث عندما يعطى المعلم مسألة تشبه تماماً مسألة محلولة في الكتاب، فقط يغير أحد أو بعض الأرقام الواردة فيها.

- يؤكد بلوم أن الغرض الأساسي من معظم ما يتعلمه التلميذ في المدرسة هو توظيفه في الحياة العملية، بمعنى أن فعالية عملية التعلم تظهر من خلال تطبيق ما يتعلمه المتعلم.

أمثلة:

- أن يفرق المتعلم بين الخلايا النباتية والخلايا الحيوانية.
- أن يحسب المتعلم درجة الحرارة الفهرنهيتية عند معرفته لدرجة الحرارة المئوية مستخدماً في ذلك العلاقة بينهما.
- أن يميز المتعلم الأحماض من بين مجموعة من المواد.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يطبق، يستخدم، يعلل، يحل (مسألة، أو مشكلة)، يحسب، يوضح.

4- مستوى التحليل Analysis:

ويُعرف وصول المتعلم إلى هذا المستوى بقدرته على تجزئة الموضوع إلى مكوناته

الأساسية أو أجزائه، بحيث يتضح التدرج الهرمي للأفكار الرئيسة فيه، وتتضح العلاقات بين هذه الأفكار والارتباط بينها (اي أن تفكير الطالب هنا ينتقل من إدراك الكليات إلى إدراك الجزئيات). وتتضمن القدرة على التحليل ثلاثة مستويات:

- تحليل العناصر.
- تحليل العلاقات.
- تحليل المبادئ التنظيمية.

أمثلة:

- أن يصنف المتعلم المواد (العناصر) إلى فلزات ولافلزات بعد معرفته للخواص العامة لكل منها.
- أن يعدد المتعلم خصائص الحديد بعد معرفته لخصائص المعادن.
- أن يستخلص المتعلم الأساس الذي بُنيت عليه تسمية المركبات الكيميائية وهي: الصودا الكاوية، الماء الملكي.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يبرهن على صحة، يقارن، يميز أو يحدد (العوامل الأساسية في ظاهرة أو موضوع)، يحلل موضوعاً إلى عناصره، يستنتج، يتعرف على.

5- مستوى التركيب Synthesis:

يصبح المتعلم قادراً في هذا المستوى على جمع عناصر وأجزاء لبناء نظام متكامل أو وحدة جديدة، فمن معلومات أو عناصر يرتبها ويربط بينها، يتوصل إلى تركيب لم يكن موجوداً قبلاً، (اي أن تفكير الطالب هنا ينتقل من إدراك الجزئيات إلى إدراك الكليات). وتتضمن القدرة على التركيب ثلاثة مستويات:

- إنتاج وسيلة اتصال فريدة.
- إنتاج خطة أو مجموعة مقترحة من العمليات.
- اشتقاق مجموعة من العلاقات المجردة

يدخل في هذا التعبير والكتابة عن المشاعر، أو التوصل إلى خطة للعمل (مثل: التخطيط لوحدة دراسية)، أو كتابة قصة حول موضوع معين أو إنتاج شعري.

أمثلة:

- أن يتوصل المتعلم إلى السلسلة الغذائية في النظام البيئي التالي:
- أن يتابع المتعلم المقالات الصحفية العلمية، ويُعد مقالاً بذلك.
- أن يستخلص المتعلم خصائص المعادن من دراسته لعدد من المعادن باستخدام التجريب المعملي.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يصمم (تجربة مثلاً)، يركب، يخطط، يقترح (أسلوباً أو طريقة)، يجمع بين، يشتق، ينظم، يعيد ترتيب.

6- مستوى التقويم Evaluation:

وهو الحكم الكمي والكيفي على موضوع، أو طريقة، في ضوء معايير يضعها المتعلم أو تُعطى له، ويشمل هذا إصدار الأحكام في ضوء معايير داخلية مثل التناغم وعدم وجود تعارض، أو في ضوء معايير خارجية (أي خارجة عن الموضوع نفسه) مثل عدم معارضة قيم وعقيدة المجتمع.

أمثلة:

- أن يعطي المتعلم رأيه في أنواع الأطعمة التي يتناولها من حيث أهميتها لجسم الإنسان.
- أن يميز المتعلم جوانب الضعف في العمود البسيط.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يصدر حكماً على، ينقد، يناقش بالحجة، يقوّم، يقدّر قيمة، يبين التناقض، يدعم بالحجة، يبرر.

ملاحظة هامة: واضح من تصنيف بلوم وزملائه للمجال المعرفي أن المعيار في التدرج الذي استخدم فيه هو درجة تعقيد العمليات العقلية، فالمستويات الدنيا (التذكر) لا تتطلب إلا قدراً يسيراً من الفهم، أو المعالجة الذهنية، بينما المستويات

العليا (التحليل - التركيب - التقويم) تتطلب أعلى درجات الفهم والقدرة على مناقشة الأفكار وتحليلها، والحكم عليها.

يمكن لتصنيف بلوم للمجال المعرفي أن يستخدم كأساس لما يلي:

1- صياغة أهداف التعليم والتعلم.
2- التأكيد على المستويات العقلية العليا.
3- إعداد خطة الدرس.

ثانيا: المجال النفسحركي (المهاري) Psychomotor Domain

هو المجال المهاري، خاصة ما يتصل بتنمية الجوانب الجسمية الحركية والتنسيق بين الحركات. كما في الأنشطة الرياضية للقيام بأداء معين. وفي هذا المجال لا يوجد تصنيف متفق عليه بشكل واسع كما هو الحال في تصنيف الأهداف المعرفية.

والأهداف التعليمية في المجال النفسحركي تصاحب الخصائص المعرفية والوجدانية لكن الخصائص النفسحركية تغلب على استجابات المتعلمين.

ويتكون هذا المجال من المستويات الآتية:

1- الاستقبال: وهو يتضمن عملية الإدراك الحسي والإحساس العضوي التي تؤدي إلى النشاط الحركي.
2- التهيئة: وهو الاستعداد والتهيئة الفعلية لأداء سلوك معين.
3- الاستجابة الموجهة: ويتصل هذا المستوى بالتقليد والمحاولة والخطأ في ضوء معيار أو حكم أو محك معين.
4- الاستجابة الميكانيكية: وهو مستوى خاص بالأداء بعد تعلم المهارة بثقة وبراعة.
5- الاستجابة المركبة: وهو يتضمن الأداء للمهارات المركبة بدقة وسرعة.
6- التكييف: وهو مستوى خاص بالمهارات التي يطورها الفرد ويقدم نماذج مختلفة لها تبعاً للموقف الذي يواجهه.

7- التنظيم والابتكار: وهو مستوى يرتبط بعملية الإبداع والتنظيم والتطوير لمهارات حركية جديدة.

تصنيف زايس: Zais

- أورد زايس Zais تصنيفاً مبسطاً لهذا المجال، يبدو مفيداً لصياغة الأهداف التعليمية النفسحركية، وجعله من أربع مستويات هي:

1- ملاحظة أداء شخص ماهر.
2- تقليد العناصر الأساسية للمهارة.
3- التمرين، ويتضمن تكرار تتابع عناصر المهارة، مع تقليل الجهد المبذول في الأداء تدريجياً.
4- إتقان المهارة مع احتمال زيادة هذا الإتقان.

تصنيف أنيتا هارو Anita Harrow:

- يقسم هذا التصنيف المجال النفسحركي إلى ست مستويات هي:

1. الحركات الانعكاسية Reflex Movements:

هي حركات لا إرادية بطبيعتها، وتظهر في مرحلة مبكرة من العمر، سواء كانت منعكسات الحاجات الأساسية، أو منعكسات حفظ التوازن والحركات الدفاعية.

2. الحركات الأساسية Fundamental Movements:

تنشأ من تجمع حركات منعكسة في أنماط أساسية، ويُستعان بها في أداء الحركات الإرادية مثل المشي والجري والقفز والرفع والجذب وتناول الأشياء.

أمثلة:

- أن يدحرج المتعلم كرة لزميل له.
- أن يمسك المتعلم القلم ليكتب.

3- القدرات الحركية الحسية Perceptual Abilities:

يحتاج المتعلم أن يدعم هذه القدرات بالنضج، وخبرات التعلم المناسبة التي تصل به إلى درجة من التحمل والمرونة، تجعله يؤدي النشاطات الحسية بمهارة، مثل القدرة على التمييز حسياً بين أشياء متشابهة كقطع العملة مثلاً، أو القدرة على التنسيق بين اليد والعين للقف الكرة مثلاً.

مثال: أن يكتب المتعلم حروف الهجاء نقلاً عن نموذج يوجد أمامه.

4- القدرات الجسمية Physical Abilities:

هي الحركات التي تتصف بالقوة والرشاقة والمرونة والتحمل، ولذلك فإن الأهداف بهذا المستوى تركز على النمو الجسمي السليم للقيام بحركات في مستوى أعلى. كما أن الخبرات التعليمية هي التي تهذب القدرات الحركية الحسية والجسمية.

مثال: أن يؤدي المتعلم ست ضغطات من وضع الانبطاح.

5- الحركات الماهرة Skilled Movements:

يُتوقع من المتعلم في هذا المستوى أن يكون قادراً على القيام بمهارات حركية عالية، كأن يكون المتعلم بعد ثلاثة أشهر من تدريبه على الطباعة قادراً على طباعة 30 كلمة في الدقيقة بما لا يزيد عن 3 أخطاء، أو أن يتقن حركة الشقلبة في الجمباز.

مثال: أن يصوب المتعلم كرة إلى الهدف بدقة.

6- التعبير الحركي المبتكر Nondiscursive Communication Movements:

بعد أن يكتسب المتعلم المقدرة على أداء الحركات الماهرة، فإنه قد يستطيع الابتكار والإبداع في أداء الحركات، بحيث يضفي عليها تعبيراً وإبداعاً وجمالاً، مثل حركات المتعلم الابتكارية لتوصيل معاني للآخرين في تمثيلية درامية على المسرح المدرسي.

ملاحظة هامة: واضح من تصنيف أنيتا هارو للمجال النفسحركي أن المعيار

في التدرج هو التخلص من الحركات الزائدة، وتقليل الجهد الإرادي في العمل بحيث تتم كثير من الحركات بصورة روتينية آلية، ليولي المتعلم اهتماماً أكبر للجانب العقلي من المهارة مما يؤدي به إلى التجديد والابتكار. ويظهر هذا واضحاً في تعلم الطفل مهارات الكتابة والقراءة مثلاً:

- على المستوى الأدنى: نجد الأفعال المنعكسة والحركات الأساسية، حيث تكثر الأفعال الزائدة والتردد وعدم الثقة، وتركيز الذهن في كل خطوة من خطوات العمل، وكل حركة من حركاته.
- وعلى المستوى الأعلى: يكون الفرد قد أصبح يؤدي كثيراً من خطوات العمل بسرعة، ودون جهد واع، فيؤدي العمل بمهارة، ويبتكر ويبرع فيه.

ثالثا: المجال الوجداني (العاطفي) Affective Domain

هو المجال الذي يحوي أهدافاً تصف تغيرات في الاهتمامات والاتجاهات والميول والقيم والتقديرات. أي أن الأهداف في هذا المجال تعتمد على العواطف والانفعالات.

يعتبر المجال الوجداني من أصعب المجالات الثلاثة في التعامل معه وتنميته، ويرجع ذلك للأسباب التالية:

1- عدم وجود تعريفات إجرائية لمكونات هذا المجال تتصف بالصدق والموثوقية.

2- الحاجة إلى وقت طويل لتنمية مكونات هذا المجال.

3- نواتج التعلم موقفية، أي قد تختلف من موقف لآخر، كما أنها خادعة إذ قد يظهر المتعلم غير ما يبطن.

- وقد صنف هذا المجال كراثوهل Krathwohl وزملاؤه عام 1964 إلى خمسة مستويات تتدرج من البساطة إلى التعقيد، من الاستقبال، إلى الاستجابة، إلى الاعتزاز بقيمة، إلى بناء نظام قيمي، إلى الاتصاف بنظام قيمي.

فيما يلي المستويات الخمس للمجال الوجداني طبقا لتصنيف كراثوهل:

1- الاستقبال Receiving:

- هو مستوى الانتباه إلى الشيء أو الموضوع، بحيث يصبح المتعلم مهتماً به، ويبدأ هذا المستوى من وقت يكون على المعلم أن يجذب انتباه المتعلم، إلى موقف يكون على المتعلم نفسه أن يولي اهتماماً بموضوعه المفضل.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يلتفت إلى، يصغي إلى، يحس بـ...، يستمع إلى، يوافق على.

مثال: أن يبدي التلميذ اهتماماً نحو موضوع التلوث.

2- الاستجابة Responding:

هو مستوى الرضا والقبول، أو الرفض والنفور، وتزداد الفعالية هنا عن المستوى السابق، فهو "يفعل" شيئاً إزاء الموضوع.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يُقبل على، يبدي إعجابه بـ..، يميل إلى، يتحمس لـ..، يعبر عن تذوقه لـ..، يبتعد عن، يعاون في، يتطوع لـ..، يستجيب، يعبر عن استمتاعه بـ...

مثال: أن ينظّف التلميذ أسنانه بعد كل طعام.

3- الاعتزاز بقيمة Valuing:

هنا يرى المتعلم أن الشيء أو الموضوع أو السلوك له قيمة ويكون هذا بسبب تقدير المتعلم نفسه لهذه القيمة، ولو أن هذا التقدير هو في الحقيقة نتاج اجتماعي تقبله المتعلم تدريجياً وببطء، حتى أصبح جزءاً منه هو وتبناه، واستخدمه كما لو كان نابعاً منه هو.

يُظهر المتعلم هذا السلوك بدرجة من الاستمرارية والثبات في المواقف المناسبة،

بحيث يشتهر عنه إنه يتصف بهذا الاتجاه أو المعتقد أو القيمة، وسعيه إلى إقناع الآخرين بها، وإخلاصه وتفانيه فيها.

تصبح الأهداف في هذا المستوى المادة الأولية التي ينمو منها ضمير الفرد وتتدخل في ضبط سلوكه.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يستحسن، يبدي رغبة مستمرة في، ينمو شعوره نحو، يتحمل مسئولية، يسهم بنشاط في، يكون اتجاهاً نحو، يحترم، يُعظّم، يهاجم، يشجب، يذم، يعارض.

مثال: أن يثق المتعلم بقدرة العلم والتكنولوجيا في حل مشكلاتنا.

4- تكوين نظام قيمي Organization:

يكسب المتعلم في تفاعلاته مع مجتمعه التعليمي (النظامي، وغير النظامي) قيماً متعددة. وعندما يصل إلى درجة كافية من النضج يبدأ في بناء نظام خاص به لهذه القيم، تترتب فيه قيمه، ويتضح فيها السائد والمسيطر منها على السلوك من الأقل تأثيراً فيه.

هذا البناء القيمي للمتعلم ليس ثابتاً، بل يمكن أن يحدث فيه تعديل مع كسبه لقيم جديدة.

حبذا لو كانت القيم التي نكسبها لتلاميذنا متجانسة متوافقة داخلياً، وليس بينها تعارض، فهذا يقوده في المستقبل إلى تكوين فلسفة للحياة، ولكن كثيراً ما يكون الواقع غير هذا.. فقد تجعله قيمه حريصاً على أداء شعائر دينه، ولكن يسمح لنفسه أن يغش في عمل، أو لا يتقن مهمة أُكلت إليه.

تستخدم في صياغة أهداف هذا المستوى أفعال مثل: يؤمن بـ..، يعتقد في، يضحي في سبيل.

مثال: أن يتبنى التلميذ القيم العلمية.

5- الاتصاف بتنظيم أو مركب قيمي، والإيمان بعقيدة

Characterization by a Value Complex

القيم التي يتبناها الفرد على هذا المستوى تكون قد وجدت لها مكاناً في الهرم القيمي له، وأصبحت ضمن نظام يتوفر فيه توافق واتساق داخلي يتحكم في الفرد.

في هذا المستوى يصل الفرد بعملية التبني الذاتي للقيم إلى القمة، بحيث تلون نظرة الإنسان إلى الكون، وتكون فلسفة له في الحياة.

من الأهداف التي تنتمي لهذا المستوى:

- أن يكوّن الفرد نظاماً للسلوك ينبني على مبادئ أخلاقية تقوم على قيم معينة (الدين، أو التضحية في سبيل الوطن، أو خدمة الإنسانية.. الخ).
- أن يكوّن الفرد فلسفة متناسقة للحياة.

يتضح من التناول التفصيلي السابق للمجال الوجداني أن المبدأ التنظيمي الذي يقوم عليه تصنيف كراثوهل لهذا المجال: هو التبني الذاتي والاستيعاب الداخلي. فعلى المستوى الأدنى نجد أقل قدر من المشاركة والالتزام، وعلى المستوى الأعلى نجد أن القيم والاتجاهات تتشربها شخصية الفرد وتتحكم في أخلاقياته.

أجزاء الهدف السلوكي:

يرى روبرت ميجر في عام 1975 أن الهدف السلوكي يجب أن يحتوي على ثلاثة أجزاء هي كما يلي:

1- وصف السلوك المرغوب تحقيقه بواسطة المتعلم بعد مروره بخبرة تعليمية.
2- وصف الحد الأدنى لمستوى الأداء المقبول.
3- وصف الشروط أو الظروف التي يتم خلالها قيام المتعلم بالسلوك المطلوب.

صياغة الأهداف السلوكية

مكونات الهدف السلوكي:

حتى يُصاغ الهدف السلوكي صياغة جيدة ينبغي أن يُضمّن:

- المصدر الصريح (مثل: كتابة أو قراءة) أو المصدر المؤول (أن + الفعل).
- الفعل أو المصدر يكون سلوكياً يمكن ملاحظته، ويمكن الاستعانة بتصنيف بلوم وزملائه للأهداف التربوية.
- المتعلم (القائم بالسلوك)، حيث يكون هو الفاعل للفعل.
- مصطلح المادة الذي يتناولها الهدف السلوكي.
- الحد الأدنى لمستوى الأداء المقبول (وهذا اختياري)، ويعتمد هذا على طبيعة الهدف، والمستوى المبدئي لأداء التلميذ.

أمثلة:

من أمثلة الأهداف السلوكية التي استخدم فيها **المصدر المؤول**:

1- أن يقيس التلميذ قطر سلك معدني دون أن يتجاوز الخطأ 2 مليمتر.

2- أن يذكر التلميذ أهمية قانون حساب سرعة الأجسام المتحركة.

3- أن يُعرّف التلميذ المقصود بعملية التنفس.

أما إذا أريد **استخدام المصدر الصريح** في هذه الأهداف فيبدأ بكتابة عبارة مثل: أن يصبح التلميذ في نهاية دراسة الموضوع قادراً على:

1- قياس قطر سلك معدني دون أن يتجاوز الخطأ 2 مليمتر.

2- ذكر أهمية قانون حساب سرعة جسم متحرك.

3- تعريف المقصود بعملية التنفس.

الأخطاء الشائعة في صياغة الأهداف السلوكية:

1- وصف نشاط المتعلم بدلا من نشاط التعلم:

الغرض من الهدف السلوكي تحديد سلوك التلميذ عند نهاية الحصة، وليس الدور الذي ينبغي أن يقوم به المعلم خلال الحصة.

مثال: (أي الهدفين التاليين أدق؟)

أ ـ أن أقوم بإثبات قانون الانكسار.

ب ـ أن يفسر الطالب رسماً يبين العلاقة بين زاوية السقوط وزاوية الانكسار.

2- تحديد موضوعات العملية التعليمية بدلا من نتائجها

يقوم بعض المعلمين بتحديد موضوع الدرس بدلاً من تحديد السلوك المتوقع من الطالب في نهاية الحصة

مثال: (أي الهدفين التاليين أدق؟)

أ- أن نقوم بدراسة قانون شارل.

ب- أن يستنتج الطالب قانون شارل.

3- وصف العملية التعليمية بدلا من نتائجها:

يركز بعض المعلمين على العملية التعليمية بدلاً من التركيز على نتائجها.

مثال: (أي الهدفين التاليين أدق؟)

أ- أن يتم تحقيق معرفة الطالب لتركيب الورقة.

ب- أن يصف الطالب تركيب الورقة.

4- صياغة أهداف مركبة:

تقوم فكرة الأهداف السلوكية على تحديد هدف لكل عبارة، لكن بعض المعلمين يقومون بصياغة أهداف مركبة.

مثال: (أي الهدفين التاليين أدق؟)

أ- أن يوضح الطالب تركيب الزهرة في النبات ووظيفتها.

ب- أن يوضح الطالب تركيب الزهرة في النبات.

اعتبارات ينبغي مراعاتها عند الأخذ بالأهداف السلوكية:

1- من الصعب صياغة نتائج العملية التعليمية كلها بصورة قابلة للقياس، وخاصة عند تنمية الأهداف في المجال الوجداني.

2- التقيد الشديد بالأهداف يضعف قابلية المعلم والطالب للتجديد والابتكار، طالما أنهم يسعون لتحقيق أهداف مخطط لها مسبقاً. فالمعلم لا يصبو فقط لتحقيق أهداف مرسومة سلفاً، بل إنه من خلال التفاعل الصفي تتكشف أشياء أخرى لديه قد لا تكون واضحة مسبقاً.

3- اختيار الأهداف على أساس مدى قياسها قد يجر العملية التعليمية نحو أهداف بسيطة سهلة التحقيق، فالمعلم عندما يشعر أنه يقيم على أساس مدى نجاحه في تحقيق الأهداف التي يحددها قبل بدء الحصة فإنه سيميل إلى كتابة أهداف سهلة التحقيق ضماناً لنجاحه في تحقيقها.

دور الأهداف السلوكية في العملية التعليمية:

أولا: دورها في تخطيط المناهج وتطويرها:

1- تسهم في بناء المناهج التعليمية وتطويرها، واختيار الوسائل والتسهيلات والأنشطة والخبرات التعليمية المناسبة لتنفيذ المناهج.

2- تسهم في تطوير الكتب الدراسية وكتب المعلم المصاحبة لتلك الكتب.

3- تسهم في توجيه وتطوير برامج التعليم الذاتي والتعليم المبرمج وبرامج التعليم بواسطة الحاسب الآلي.

ثانيا: دورها في توجيه أنشطة التعلم والتعليم:

1- تيسر عملية التفاهم بين المعلمين من جهة وبين المعلمين وطلابهم من جهة أخرى، فالأهداف السلوكية تمكن المعلم من مناقشة زملاءه المعلمين حول الأهداف والغايات التربوية ووسائل وسبل تنفيذ الأهداف مما يفتح المجال أمام الحوار والتفكير التعاوني مما ينعكس ايجابيا على تطوير المناهج وطرق التعليم. كما أنها تسهل سبل الاتصال بين المعلم وطلابه فالطالب يعرف ما هو مطلوب منه وهذا يساعد على توجيهه وترشيد جهوده مما يساعد على التقليل من التوتر والقلق من قبل الطال حول الاختبارات.

2- تسهم الأهداف السلوكية في تسليط الضوء على المفاهيم والحقائق والمعلومات الهامة التي تكون هيكل الموضوعات الدراسية وترك التفصيلات والمعلومات غير الهامة التي قد يلجأ الطالب إلى دراستها وحفظها جهلاً منه بما هو مهم وما هو اقل أهمية.

3- توفر إطارا تنظيميا ييسر عملية استقبال المعلومات الجديدة من قبل الطالب فتصبح المادة مترابطة وذات معنى مما يساعد على تذكرها.

4- تساعد على تفريد التعلم والتعامل مع الطالب كفرد له خصائصه وتميزه عن غيره من خلال تصميم وتطوير برامج التعلم الذاتي الموجهة بالأهداف والتي يمكن أن تصمم في ضوء مجال خبرات الطالب واستعداده الدراسي.

5- تساعد على تخطيط وتوجيه عملية التعلم عن طريق اختيار الأنشطة المناسبة المطلوبة لتحقيق التعلم بنجاح بما في ذلك اختيار طريقة التدريس الفاعلة والمناسبة للأهداف واختيار وسائل التعليم المفيدة لتحقيق الهدف السلوكي.

6- تساعد المعلم على إيجاد نوع من التوازن بين مجالات الأهداف السلوكية ومستويات كل مجال من المجالات.

7- توفر الأساس السليم لتقويم تحصيل الطالب وتصميم الاختبارات واختيار أدوات

التقويم المناسبة وتحديد مستويات الأداء المرغوبة والشروط أو الظروف التي يتم خلالها قياس مخرجات التعلم.

8- ترشيد جهود المعلم وتركيزها على مخرجات التعلم (الأهداف) المطلوب تحقيقها.

9- تعتبر الأهداف السلوكية الأساس الذي تبنى عليه عملية التصميم التعليمي ونتاج هذه العملية عبارة عن نظام يلائم المتغيرات في الموقف التعليمي.

10- تيسر التفاهم والاتصال بين المدرسة ممثلة بمعلميها وهيئتها التدريسية وبين أولياء الأمور فيما يتعلق بما تود المدرسة تحقيقه في سلوك الطالب نتيجة للأنشطة المتنوعة التي تقدمها لهم في المجالات المختلفة (معرفية، نفسحركية، وجدانية).

ثالثا: دور الأهداف في عملية التقويم:

تقوم الأهداف على توفير القاعدة التي يجب أن تنطلق منها العملية التقويمية. فالأهداف تسمح للمعلم والمربين بالوقوف على مدى فعالية التعليم ونجاحه في تحقيق التغير المطلوب في سلوك المتعلم ما لم يحدد نوع هذا التغير أي ما لم توضع الأهداف فلن يتمكن المعلم من القيام بعملية التقويم مما يؤدي إلى الحيلولة دون التعرف على مصير الجهد المبذول في عملية التعليم سواء كان هذا الجهد من جانب المعلم أو المتعلم أو السلطات التربوية الأخرى ذات العلاقة.

الفصل الثالث

خطوات بناء الاختبار أو المقياس

يحتوي الفصل على:

- تحديد الهدف أو الأهداف الرئيسة للاختبار.
- تقرير محك أو معيار الدرجة.
- ترجمة المفاهيم والأهداف إلى خصائص محددة.
- تصميم فقرات مناسبة تعبر عن هذه الخصائص.
- اختبار مستوى الصدق والصعوبة للفقرات.
- فعالية المشتتات.
- إعداد الاختبار للاستخدام.
- تقنين الاختبار.

خطوات بناء الاختبار أو المقياس:

يخضع تصميم وإعداد الاختبار الجيد لعدد من الاعتبارات والقواعد المنهجية، كما تتطلب فنية تصميم الاختبار دراية واسعة وعميقة بالأساليب المختلفة ومدى مناسبة كل منهما للأغراض المختلفة. وأساليب تصميم وانتخاب فقرات اختبار

صادقة وثابتة ومميزة هي أساليب عملية وعلمية معاً. ويتضمن الجانب العلمي فيها المعرفة الدقيقة بكيفية وضع الفقرة ومهارة تصميمه مع حساب قدرته التمييزية ومستوى صعوبته وثباته ويأتي الجانب الفني في الربط التكاملي بين هذه الخصائص العلمية والخصائص اللفظية للفقرة والاعتبارات العلمية التي سيستخدم فيها الاختبار ولا يعد تصميم وأعداد اختبار جديد هدفاً في حد ذاته، فالباحث لا يلجأ عادة لهذا العمل إلا إذا دعته إلى ذلك حاجة ملحة وعندما يتضح له أن الاختبارات المتوفرة لا تفي بغرضه، ودون وجود الدواعي الضرورية لتصميم اختبار جديد، يظل استخدام الاختبارات المتوفرة في المجال أو تطويرها أفضل كثيراً للاستفادة من الدراسات والنتائج التي سبق التوصل إليها.

ويمكننا أن نلاحظ أن هناك اتجاهين بارزين في تصميم الاختبارات وهي:

الاتجاه الأول: علمي نظري يهدف إلى تصميم الاختبار وفق إطار نظري محدود، وللإجابة على فروض جيدة الصياغة يقدمها العالم. ويصمم وينتخب بعناية فائقة فقرات الاختبار وفقاً لمدى قربها من فروضه الأساسية.

الاتجاه الثاني: عملي أو فني وهو يرمي لمجرد تطوير اختبارات جديدة لتكون أداة في يد الممارس أو الأخصائي النفسي بغض النظر عن اعتبارها أداة علمية للبحث.

وَالأختلاف بين هذين الاتجاهين له تأثيرات هامة على منهجية تصميم الاختبارات، فمصححوا الاختيارات ممن يأخذون بالمنحنى العملي يطورون أساليب خاصة بهم تختلف بقدر ما عن تلك التي يستخدمها العالم الذي يصب القدر الأكبر من اهتمامه على صياغة فروضه النظرية. ورغم أن صياغة الفروض جوهرية للغاية في تصميم الاختبار، إلا أن عمليات صياغة هذه الفروض تختلف تماماً في الحالتين بين العالم والفني فبالنسبة للعالم تعد مشكلة صياغة الفروض أهم خطوات العمل كلها وأصعبها وهي ما يتطلب منه الأصالة إذ لابد أن تكون فروضه متجهة إلى تقديم إضافة حقيقية مع اتساقها في الوقت نفسه مع الحقائق القائمة والملاحظات

السابقة والتبريرات المنطقية التي يقدمها العالم لفروضه هي التي تميز عمله عن عمل الفني الذي يهدف إلى هدف محدود ويحكم إجراءاته المهارة في أصابه الهدف وتحقيق الأغراض المباشرة ورغم ذلك نستطيع أن نجد المزايا والعيوب في كلا المنحنين فالمنحى العملي الفني سريع النتائج، وتصميم الاختبارات فيه لا يتطلب وقتاً طويلاً، وقد يبدأ من معلومات خاطئة ولكنه يؤدي لنتائج عملية مفيدة ينتهي بمعلومات محدودة ذات قيمة نظرية ضئيلة. بينما المنحى العملي النظري يتطلب وقتاً في إعداد الاختبارات، ووقتاً في تنظيم وتحليل النتائج وقد تكون نتائجه العملية محدودة ولكنه يبدأ من فروض صلبة وينتهي بمعارف هامة ويؤدي في أحيان كثيرة إلى الإسهام في تطوير النظرية.

خطة تصميم الاختبار:

تقسم خطة تصميم الاختبار إلى عدد من الخطوات المحددة وتختلف هذه الخطوات وتصنيفها من باحث لأخر.

إلا أنها تتضمن في كل الحالات في البداية التي تبدأ بها وتنتهي إليها كالأتي:

1- تحديد الهدف أو الأهداف الرئيسة من الاختيار.

2- تقرير محك أو معيار الدرجة.

3- ترجمة المفاهيم والأهداف إلى خصائص محددة.

4- تصميم فقرات مناسبة تعبر عن هذه الخصائص.

5- اختبار مستوى الصعوبة والصدق للفقرات.

6- فعالية المشتتات (المموهات)

7- إعداد الاختبار للاستخدام.

8- تقنين الاختبار.

تحديد الهدف أو الأهداف الرئيسة من الاختبار

إن تحديد الهدف من الاختبار من حيث ميادين استخدامه والمجتمع المراد تطبيقه عليه، على جانب كبير من الأهمية. فإن كان الهدف من استخدام الاختبار في الحصول على بيانات دقيقة كان لابد من استخدام الطرق الإحصائية والفنية التي تحقق هذا الغرض، أما إذا كان المطلوب جمع بيانات سريعة أو إذا كان المفروض أن يعهد بالاختبار إلى فئة من المعلمين غير المدربين تدريبياً كافياً على استخدام الاختبارات المقننة روعي في تصميم الاختبار سهولة تفسير معاييره حتى ولو أدت هذه السهولة إلى التضحية ببعض الشيء ببعض الدقة ذلك لأن تعقيد الاختبار من ناحية التعليمات أو التصحيح سيؤدي حتماً إلى أخطاء أكبر.

إذن رغم أننا قد نبدأ من مفهوم علمي مقبول يمكن اختباره بواسطة اختبار معين إلا إن تحديد الهدف من الاختبار يتطلب مزيداً من التفصيل فقد يكون الهدف من اختبار ما تحديد مستويات الأفراد وفقاً لمقدار امتلاكهم للخاصية وقد يكون الهدف من اختبار آخر التمييز بين الأفراد وفقاً لترتيبهم على الخاصية أو قد يتكون الهدف تحديد الصعوبات أو الاضطرابات التي يعاني منها الفرد وبالتالي تصبح *مهمة الاختبار تحليل هذه الصعوبات بشكل تفصيلي وتصنيفي.*

ويختلف كل هدف من هذه الأهداف في طبيعته، فالأساليب التي تهدف إلى تحديد مستوى الأفراد يكثر استخدامها في المجالات التحصيلية أو في قياس القدرات المعرفية. مما يتطلب تصنيف وتحليل العمليات المختلفة في المادة المحصلة وحيث يمكن استخدام هذا التصنيف كدليل وصفي لناتج العملية التعليمية ومدى تحقق أهدافها.

تقرير محك أو معيار الدرجة:

تفرض أهداف الاختبار نوع الإطار المرجعي الذي تفسر في ضوئه درجة الفرد وما إذا كان يعد جيد الأداء أم لا ويمكن التمييز هنا بين نوعين من المحكات:

أ- المحك المرجعي: ويستخدم هذا المحك أساساً في اختبارات التحصيل فقياس التحصيل يتضمن إشارة إلى حدود مقبولة للأداء على متصل يبدأ قطبه الأدنى من لا كفاءة على الإطلاق وينتهي قطبه الأعلى بأداء محكم تماماً وفق تعريفات وحدات الاختبار مسبقة التحديد.

وحتى يمكن تصميم اختبار جديد ذي محك مرجعي فمن الضروري العمل وفق خطوات محددة من ذلك ضرورة وضع أشكال الأداء المطلوبة بصورة صريحة ومباشرة. وضرورة تحديد محكات قبول الأداء مسبقاً وفقاً لبرنامج تحصيل الموضوع المعين مثلاً ويجب تحليل الأداء المطلوب لعناصره المختلفة حتى يمكن تقدير جوانب تحصيل المادة أو الخبرة موضوع الاختبار وأخيراً فإن النجاح على الاختبار يتقرر من خلال مقارنة الأداء لا بأداء أفراد آخرين ولكن بمقارنته بحجم المطالب المحددة في الاختبار وما أنجز منها. وسيتم لاحقا تحديد الخطوات التفصيلية لبناء الاختبارات المحكية المرجع.

بـ المحك المعياري: تختلف الاختبارات ذات المحك المعياري عن الاختبارات المحكية المرجع في أن الدرجة الخاصة بالفرد لا تقارن وفق حجم سابق التحديد للمطلب موضوع الاختبار، وإنما تقارن بالأداء الخاص ببقية الأفراد، فنحن هنا نهدف للتعرف على الوضع النسبي للفرد بين مجموعة الأفراد، وبهذا تكون الدرجة الخاصة بالفرد ذات قيمة من حيث أنها تفيدنا بمدى جودة هذا الأداء بالنسبة للمجموعة المختبرة ولكنها ليست ذات فائدة من حيث الإشارة إلى مدى جودة الأداء بالنسبة للمطالب الأساسية للاختبار أو المجال موضوع القياس. ويترتب على تقرير هذا المحك المعياري أن نتجه في تصميمنا للاختبارات إلى أن تكون الفقرات ذات قدرة على تفسير مدى الدرجات التي يمكن أن نحصل عليها من أية مجموعة من الأفراد.

معنى هذا أن الدرجة على المقياس في حد ذاته لا معنى لها على الإطلاق إلا إذا نسبت إلى إحدى المحكين: المحك المرجعي بأن تنسب إلى الأداء نفسه بوصفه

مرجعاً للدرجة أو المحك المعياري بأن تنسب إلى درجات بقية الأفراد بوصفهم معياراً لهذه الدرجة.

ترجمة المفاهيم والأهداف إلى خصائص محددة:

المفاهيم العلمية كالذكاء والانبساط والتطرف والاجتماعية تجريدات لخصائص مفترضة في الأشياء ونستدل على هذه الخصائص من وقائع سلوكية محددة سواء في شكل أفكار معبر عنها أو حلول لمشكلات أو استجابة لمنبهات محددة أو أساليب تعامل مع البيئة أو صفات شخصية تتسم بالاستقرار.

ونقوم بترجمة هذه المفاهيم إلى خصائص محددة بصورة تسمح بصياغتها في وحدات معيارية للقياس ففي مفهوم مثل الطلاقة يمكن ترجمته إلى أعمال محددة إذا وضع له التعريف الإجرائي المناسب مثل "الطلاقة هي القدرة على إنتاج أكبر عدد من الأفكار ذات الدلالة استجابة لمنبه معين".

إذ يتيح هذا التعريف الإجرائي وضع المنبهات لاستدعاء استجابات معينة هي التي يقاس بواسطة كمها واستيفائها لمحك الدلالة لطلاقة المفحوص، ولم يكن متاحاً بالمثل قياس مفهوم الذكاء دون توفر التعريف الإجرائي الذي وضعه بينيه وترجمه إلى أعمال محددة يقوم بها الطفل. فالمطلوب في هذه الخطوة هو أن نحول الأهداف إلى سلسلة من الخطوات والأعمال لنتمكن من كتابة الفقرات المناسبة التي يمكن أن يقدم تقديراً صادقاً للجوانب التي تقوم بقياسها ويتم ذلك بتحديد عينه مقننة من هذا السلوك والذي يتعين أن يمثل بشكل جيد المفهوم الذي نصمم من أجله المقياس من خلال خطوات محددة وهي:

أ- وصف السلوك المطلوب قياسه:

فإذا كان المفهوم هنا هو "القدرة الحسابية» فإن وصف السلوك يكون كالأتي: يتضمن السلوك المطلوب قياسه القيام بحل عدد من المشكلات الحسابية البسيطة

دون أخطاء ويمكن إجراء العمليات الحسابية جميعها أما داخلياً ((عقلياً)) أو باستخدام الورقة والقلم ويقصد (بالعمليات الحسابية البسيطة) الجمع والطرح والضرب والقسمة والنسب المئوية والكسور. ولأهمية السرعة في المطلوب والإجابة الصحيحة هي المطلوبة ولذلك يمنح المفحوص وقتاً متسعاً للإجابة.

ب ـ تحليل السلوك المطلوب قياسه:

ويقصد به تصنيف هذا السلوك للتعرف على مكانه بين أشكال السلوك الأخرى والإشارة إلى طبيعته والى عموميته وقابليته للتنبؤ، فإذا طبقنا مثل هذا التحليل على "القدرة الحسابية" فيمكن أن نصل إلى أن القيام بعمليات حسابية بسيطة يتأثر بكل من الاستعداد والتدريب أو العادات المستقرة لفترة طويلة والأفراد البارزين في هذا المجال لا يحتاجون لبذل مجهود كبير لتجنب الأخطاء، وحساباتهم دائماً صحيحة بصورة طبيعية، والأفراد الذين يفتقدون هذه الخاصية يقعون بالخطأ بسهولة وإجاباتهم الصحيحة تعود في حقيقة الأمر إلى مراجعتهم الدقيقة وإعادة المراجعة، والواقع أن الأشخاص أصحاب الدرجة المرتفعة يتمتعون بما يمكن اعتباره القدرة على المراجعة الآلية

ج. تقرير خصائص الفقرات:

ويقصد بها التقدم نحو وصف نوع الفقرات التي يمكن أن تؤدي إلى صدق قياس وتقويم السلوك المعين فإذا تقدمنا من المثال الخاص بالقدرة الحسابية فسنجد أن قياس السلوك الحسابي يتطلب أن يتضمن الاختبار عدداً كبيراً من المشكلات البسيطة نسبياً والتي تتضمن عمليات جمع وطرح وضرب وقسمة وكسور عشرية ونسب ويتعين أن لا تتضمن الفقرات مشكلات صعبة أو بها خدعة يجب معرفتها للوصول إلى الحل ويجب أن تكون التعليمات الموجهة للمفحوص واضحة ومتضمنة وأن يعمل المفحوص حسب سرعته الخاصة، ولكن دون أنفاق وقت كبير في المشكلة أو فقرة ما أكبر مما يجب.

تصميم فقرات مناسبة تعبر عن هذه الخصائص:

أن هناك أكثر من شكل وأسلوب يمكن أن تصاغ به فقرات الاختبار. وتتطلب كتابة فقرات مناسبة لاختبار جديد، تحليل الفقرات تحليلاً كيفياً من حيث شكل الفقرات ومن حيث مضمونها إلى فقرات جيدة تقيس بشكل صادق السلوك المعين كما يؤدي هذا التحليل الكيفي إلى توضيح أهمية مراعاة شروط كتابة الفقرات الجيدة.

كما يتطلب الأمر تحليلاً كمياً يهدف لتقدير مستوى صعوبة الفقرات وانتخاب الفقرات ذات الصعوبة المناسبة لأهداف الاختبار. ويوضح هذا النوع من التحليل علاقة صعوبة الفقرات بثباتها، إذ أن صدق وثبات الاختبار يعتمد أساساً على خصائص الفقرات ويمكن التوصل إلى صدق وثبات مرتفعين للاختبار منذ البداية من خلال إجراءات تصميم الفقرات.

التحليل الكيفي:

تصنف الاختبارات من حيث بنيتها إلى اختبارات محددة البنية أو الاختبارات الموضوعية التي تتميز بأن المفحوص لا يقدم فيها أجابه للفقرة بل ينتخب إحدى الإجابات من بين عدد من الاختيارات التي يقدمها الاختبار، وتعتبر موضوعية حيث أن تصحيح الاختبار يتم بدون تدخل من جانب الفاحص. كما أن لهذه الاختبارات القائمة على أساس الاختيار أنواع نذكر منها:

1- اختبارات الصواب والخطأ.
2- اختبارات المطابقة.
3- اختبارات الاختيار من المتعدد.

وهناك نوع آخر من الاختبارات والتي تتضمن فقرات غامضة أو لا تستدعي استجابات موضوعية التصحيح بل تحتمل تأويلات وتفسيرات مختلفة قد تختلف من باحث لأخر أو وفق مبدأ للتصحيح أو آخر، وهذه الاختبارات تسمى الاختبارات الأسقاطية.

وفي فئة ثالثة نجد الاختبارات الإنتاجية والتي تتطلب من المفحوص ((تقديم)) استجابات قابلة لتصنيفات واضحة المعالم محددة بشكل مسبق مثلما نجد في اختبارات الإبداعية أو بعض المقاييس الفردية للذكاء كاختبار الفهم أو المعلومات.

وقبل كتابة فقرات الاختبار على واضع الاختبارات أن يفكر في طول الاختبار وعدد فقراته، وأن أهداف الاختبار هي التي تحدد طوله فإن كان لأجل مقارنة سريعة فلا داعي للاختبار الطويل. أما إذا كان لأجل جمع بيانات دقيقة فمن الواجب أن يكون الاختبار طويل ومتعدد الأسئلة.

وهناك عدد من القواعد العامة في كتابة الفقرات العامة في الوصول إلى مستوى صدق جيد للاختبار منها ما يلي:

1- احتواء الفقرة الواحدة على فكرة واحدة فقط والابتعاد عن الأسئلة المزدوجة.
2- الابتعاد عن استخدام الكلمات الغريبة أو الكلمات التي تحمل أكثر من معنى واحد.
3- محاولة استخدام الصيغة الإيجابية للفقرات كلما أمكن ذلك.
4- تجنب استخدام الإطلاق في العبارات مثل دائماً ابداً، حتماً وفي كل مكان.
5- البساطة والسهولة وعدم الغموض.
6- الموضوعية وإمكانية تبويب الإجابات.
7- تحاشي الأسئلة الإيحائية.
8- تحاشي الأسئلة التطفلية والتي تسبب إحراجاً.

يضاف إلى ذلك عدد من القواعد الأساسية في تصميم الفقرات تتعلق بصميم نظرية القياس والأسس الرياضية لمعنى الدرجة على المقياس. فمن الضروري أن يسعى مصمم الاختبار لمحاولة الاقتراب منها إذ يبدو بعضها عسيراً في حالات كثيرة لهذا يحسن اعتبارها مؤشرات أساسية لتصميم الفقرات الجيدة للاختبار وتتلخص هذه القواعد في الآتي:

1- أن يتضمن الاختبار فقرات تشكل بصورة جيدة عينة ممثلة للمجال الذي تهدف لقياسه. أي أن تكون هذه العينة للسلوك الذي سبق أن قمنا بتحديده من خلال تحديد الهدف من الاختبار.

2- أن تتدرج فقرات الاختبار وفقاً لمحك معين. وتعني هذه الخاصية أن ترتب فقرات الاختبار من حيث الصعوبة بشكل متدرج وبحيث يكون مستوى صعوبتها موحداً بالنسبة لكل الأفراد. فإذا كان للفقرات مثل هذا الترتيب المتدرج فإن أي مفحوص سيتمكن من تقديم أجابه صحيحة لعدد من الفقرات عند حد معين لا يستطيع تجاوزه يقابل مستوى الصعوبة الملائم له، بحيث تعني الدرجة 17 مثلاً بالنسبة للفرد إنه أجاب أجابه صحيحة على أول 17 فقرة وأجاب أجابه خاطئة على بقية الاختبار. ويؤدي مثل هذا الترتيب للفقرات إلى إلغاء الغموض في تفسير الدرجة على الاختبار بحيث تعني الدرجة 17 نفس المعنى لدى كل من حصل عليها وهو أنهم جميعاً أجابوا أجابه صحيحة على الـ 17 فقرة الأولى.

3- أن ترتيب فقرات الاختبار وفق متصل المتغير موضوع القياس بطريقة تجعلنا أمام مقياس مسافات ويلاحظ هنا أن الفقرات التي يتضمنها مثل هذا المقياس تتسم بخاصية التعبير عن فروق متساوية على القدرة. بين الفقرات المتجاورة فيعبر الفرق بين الدرجتين 54، 53 عن فرق مماثل بين الدرجتين 86، 85 وهكذا.

4- أن يكون الصفر على أول فقرات المقياس تعبيراً جيداً عن كمية صفرية من القدرة التي تقاس فإذا أضفنا الخاصية السابقة لهذه الخاصية يصبح مقياسنا مقياس نسبة يمكن معالجة نتائجه بالعمليات الحسابية الأربعة.

5- أن تكون الفقرات بمثابة وحدات قياس ذات معنى واضح، لأن ذلك يؤدي إلى ربط وحدات القياس بالخطأ المعياري للمقياس، وهو ما يوفر ميزة تسمح لمستخدم الاختبار بترجيح أن الفروق بين الأفراد تعود أساساً إلى الخطأ في الأداء وليس لأي عامل آخر.

تقودنا كل هذه الاعتبارات والقواعد العامة والخاصة إلى الانتقال للجانب الأساسي التالي في تصميم الفقرات وهو التحليل الكمي الذي يتعلق بدراسة وتحديد وترتيب مستوى صعوبة الفقرات في الاختبار.

اختبار مستوى الصدق والصعوبة لفقرات الاختبار

التحليل الكمي:

هناك اعتباران أوليان يبرزان عند انتخاب فقرات اختبار ما.

الأول هو: هل الفقرة صادقة في قياس السمة التي نرغب في قياسها بمعنى هل تمكننا هذه الفقرة من التمييز بين الأشخاص أصحاب القدر الضئيل منها؟ ويجاب عادة على هذا السؤال من خلال الأساليب الإحصائية الخاصة بحساب صدق فقرات الاختبار.

أما **الاعتبار الثاني** فهو هل مستوى صعوبة الفقرة مناسبة للمجموعة من الأفراد الذين سيختبرون بهذا الاختبار؟

1- صدق الفقرات:

ويعني مؤشر الصدق حسن قياس الفقرة أو تمييزها في أتساق مع بقية الاختبار أو حسن تنبؤه بالمحك الخارجي. ويمكن أن نلاحظ عدداً من المؤشرات الشائعة للصدق في الاستخدامات المختلفة من ذلك: نسبة المفحوصين الذين يجيبون على الفقرة أجابه مقبولة (تمييزية) أو الارتباط بين الفقرة ومحك خارجي أو الارتباط بين الفقرة وبقية الاختبار وهو محك داخلي.

ونادراً ما يلجأ مصمم الاختبار إلى حساب الارتباط بين الفقرة والمحك الخارجي وقليلاً ما يحسب الارتباط بين الفقرة وبقية الاختبار والأسلوب الأكثر شيوعاً هو حساب القدرة التمييزية للاختبار من خلال نسبة المفحوصين الذين يتجاوزونه

ودون إجراء مثل هذا التحليل لصدق الفقرات. ومهما كانت خبرة مصمم الاختبار وموضوعيته فسيظل الاختبار في حاجة إلى فحصه لتقرير هذه الخصائص التمييزية للفقرات.

تستخدم المجموعات المتطرفة أو المتعارضة عادة في حساب صدق الفقرات وذلك بأن ننتخب مثلاً الفقرات التي يجيب عليها أعلى 10 ٪ من المجموعة (أ) وأقل 10 ٪ من المجموعة (ب) المتطرفة معها في السمة المقاسة وبقدر تعارض المجموعتين بقدر حدة التمييز الذي نتوصل إليه. إلا أن استخدام المجموعات الشديدة التطرف بهذه الصورة يؤدي إلى خفض الثبات نتيجة للعدد الصغير من الحالات المستخدمة حيث تكون المجموعتان أ ، ب عبارة عن الحالات المتطرفة في عينة واحدة مستخدمة لحساب صدق الفقرات. ويقترح كيلي Kelly استخدام أعلى وأدنى 27٪ من التوزيع باعتبارهما المجموعتين المتطرفتين بشرط اعتدالية التوزيع. بينما يرى كيرتون cureton أن النسبة مقبولة إذا كانت تتراوح بين 27٪ - 33٪ وإذا كان المنحنى أكثر استواءً من المنحنى الاعتدالي التقليدي. وتضيف أنستازي إلى أن خطأ العينة يصبح كبيراً في حالة العينات الصغيرة. ولهذا يصبح من الأفضل أن لا نحدد هذه النسب بشكل حاسم وتعتبر النسب التي تتراوح بين 5 ٪، 33٪ وافية بالغرض بصورة مقبولة، ومؤشر التمييز المقبول بصفة عامة باعتباره مؤشراً لصدق الفقرات هو ما توضحه **المعادلة الآتية لجونسون Johnson**:

يتم حساب معامل التمييز بهذه الطريقة بطرح عدد الأفراد الذين اختاروا الخيار الصحيح للسؤال في المجموعة الدنيا من عدد الأفراد الذين اختاروا الخيار الصحيح للسؤال من المجوعة العليا ثم قسمة الناتج على عدد الأفراد في احدى المجموعتين فقط.

$$م_{ع\,د} = \frac{أ_{ع} - أ_{د}}{ن2}$$

م ع د = مؤشر الفرق بين أعلى وادنى المجموعتين

أ ع = عدد الإجابات الصحيحة على الفقرة في المجموعة العليا

أد = عدد الإجابات الصحيحة على الفقرة في المجموعة الدنيا
ن = عدد أفراد العينة في إحدى الفئتين

كما يمكن إيجاد تباين كل فقرة حسب المعادلة الآتية:

تباين الفقرة = نسبة الأفراد الذين أجابوا إجابة صحيحة عن الفقرة × نسبة الأفراد الذين أجابوا إجابة خاطئة عن الفقرة

ولكن يعاب على هذه الطريقة أن التوزيع العيني لقيمتها غير معروف وبالتالي لا نستطيع أن نجيب عن الأسئلة التالية:

- هل قيمة معامل التمييز الناتجة اكبر من الصفر؟
- ما حجم الفرق بين قيمة معامل التمييز والصفر؟
- هل الفرق دال إحصائيا؟

ورغم ذلك استمرت هذه الطريقة تستخدم كواحدة من الطرق العملية لحساب معامل التمييز وتصلح لاختبارات المدرس لسهولة استخدامها وحسابها وتفسيرها.

ويمكن حساب معامل التمييز باستخدام معامل الارتباط المنصف للسلسلة ويتم استخدام هذه الطريقة عندما يكون لدينا متغيران أحدهما ثنائي الدرجة. درجة كل فقرة: (صفر أو واحد) والآخر فئوي متصل (الدرجة الكلية) حيث يعتبر معامل الارتباط المنصف للسلسلة أكثر مناسبة للاستخدام بين سؤال موضوعي والدرجة الكلية على الاختبار أو على محك خارجي مثل تقديرات المعلمين أو درجات الطلاب في اختبارات أخرى.

ويمكن حساب معامل التمييز باستخدام معامل الارتباط المنصف للسلسلة الحقيقي وتستخدم هذه الطريقة عند تقسيم المتغير المتصل (الدرجات الكلية للاختبار) إلى قسمين متساويين يشتمل أحدهما على الدرجات العليا والآخر على الدرجات الدنيا، حيث ينتج لدينا متغيران أحدهما ثنائي وهو (درجة كل فقرة) والآخر فئوي متصل (الدرجات الكلية) ولكن تم تقسيمه إلى متغير ثنائي.

$$ر\ ب\ س = \frac{م\ س - م\ ك}{ع\ س} \times \left(\frac{ص}{أ} \right)$$

ر ب س = معامل الارتباط المنصف للسلسلة

م س = متوسط الدرجة الكلية في الاختبار للمجموعة التي إجاباتها صحيحة على الفقرة

م ك = متوسط الدرجة الكلية في الاختبار(أو المحك) لكامل المجموعة

ع س = الانحراف المعياري للدرجة الكلية في الاختبار

ص = صعوبة الفقرة

أ = ارتفاع منحنى التوزيع الاعتدالي عند الدرجة (ز) التي تقسم التوزيع عند النسبة المعيارية للصعوبة (ص)

متى تستخدم الطرق الثلاث السابقة:

أ- في حالة الأسئلة متوسطة الصعوبة لا يوجد فرق بين الطرق الثلاث وبالتالي فإن طريقة مقارنة الأطراف تكون هي الأفضل وذلك لسهولة حسابها.

ب- إذا كان الغرض اختبار مدى بعد قيمة معامل التمييز عن الصفر فتستخدم إحدى الطريقتين الأخريين.

ج- أما إذا كان الغرض اختيار أسئلة ذات مستوى صعوبة متطرفة (سهلة أو صعبة) ولكن أيضاً تميز بين الأطراف فتستخدم طريقة معامل الارتباط المنصف للسلسلة.

د- إذا كان من بنى الاختبار في شك أن العينات في المستقبل سوف تختلف في القدرة عن عينة التقنين فيستخدم معامل ارتباط ثنائي التسلسل حتى يحصل على معامل تمييز عالي.

هـ- إذا كان من بنى الاختبار على ثقة بأن العينات في المستقبل سوف لا تختلف في القدرة عن عينة التقنين فيستخدم معامل الارتباط المنصف للسلسلة الحقيقي.

و- إذا كان السؤال والمحك ثنائيا فإنه يمكن استخدام معامل الارتباط المنصف للسلسلة الحقيقي.

وكلما ارتفع معامل تمييز الفقرة كلما كان إسهامه أفضل في زيادة ثبات الاختبار ورفع قيمة تباينه وعلى هذا الأساس يمكن تقويم معاملات تمييز الفقرات بناء" على المعيار الموضح في الجدول أدناه والذي وضعه ايبل 1963

مستوى التمييز	التقويم
من 0,40 فأعلى	فقرة جيد جداً
من 0,30 إلى 0,39	فقرة جيدة بدرجة معقولة ولكن يمكن تحسينها
من 0,19الى0,29	فقرة هامشية تحتاج إلى تحسين
اقل من 0,19	فقرة ضعيفة تحذف أو تعدل

2- صعوبة الفقرات:

أن فعالية الاختبار تتأثر بشكل مباشر في عدد من المواقف بتشتت قيم صعوبة الفقرات. وبصيغة عامة تحدد صعوبة الفقرات بحساب نسبة الأفراد الذين يجيبون على الفقرة أجابه صحيحة، فإذا كانت الإجابة على الفقرة أما صواب أو خطأ فالأسلوب المباشر لتحديد صعوبته هو حساب متوسط الدرجة عليه أو نسبة الإجابة بصواب للمجموع الكلي للأفراد. وهي تمثل الاحتمالية العملية لقدرة الجمهور المعني بالفقرة أو المتميز بالقدرة على الإجابة عليه أجابه صحيحة.

ولأن إمكانية توحيد مستوى الصعوبة بشكل حاسم في كل فقرات الاختبار تؤدي إلى أثار جانبية على كل من صدق الاختبار وتجانسه الداخلي، فيتعين الاهتمام في هذه الحالة بترتيب الفقرات من حيث مستوى صعوبتها من الأسهل إلى الأصعب على أن يكون مدى الصعوبة بين أسهل وأصعب فقرة محدوداً بقدر الإمكان ويدور حول نسبة 50٪ صعوبة ويؤدي ترتيب الفقرات من حيث الصعوبة إلى عدد من المزايا:

من ذلك تنمية ثقة المفحوص بإجابته على الفقرات السهلة أولاً قبل أن يلتقي بالفقرات الصعبة قرب الحد المقابل لأقصى قدرته، كما يؤدي هذا الترتيب إلى خفض إمكانية تضييع جزء من الوقت في فقرات تفوق قدرة الفرد وإهماله لفقرات سهلة وفي متناوله.

ويقتصر تحليل مستوى الصعوبة على اختبارات الاستعدادات والقدرات والتحصيل وهو أكبر أهمية بالنسبة لاختبارات الاستعدادات والقدرات أما اختبارات التحصيل فيحدث أحياناً أن يكون تقدير خبير في الموضوع أكثر أهمية من حساب مستوى الصعوبة إحصائياً. كما أن اختبارات السرعة بكل أشكالها تخرج عن إطار مشكلة الصعوبة إذ تتضمن أساساً فقرات سهلة ويفضل اختيار الفقرات ذات الصعوبة المتوسطة وبمدى ضيق حول متوسط الصعوبة (0,50) لأنه يعطي أعلى قيمة تباين للفقرة، وهو ما يؤدي إلى ارتفاع قيمة تباين الاختبار ككل وبالتالي ارتفاع قيمة ثباته، أما في اختبارات القوة فيفضل تدرج فقرات الصعوبة من الأسهل للأصعب، وبالتالي احتواء الاختبار على مدى واسع من مستويات الصعوبة. ويمكن الوصول إلى مقياس أفضل لمستوى الصعوبة بواسطة تحويل نسب الصعوبة إلى درجات معيارية باعتبار النسب من الأفراد الذين يجيبون أجابه صحيحة على الفقرة مقابلة للنسب تحت المنحنى الاعتدالي.

$$\text{معامل الصعوبة} = \frac{\text{ص}_{\text{ع}} + \text{ص}_{\text{د}}}{2\,\text{ن}}$$

ص ع = عدد الأفراد في الفئة العليا الذين أجابوا إجابة صحيحة

ص د = عدد الأفراد في الفئة الدنيا الذين أجابوا إجابة صحيحة

ن = عدد الأفراد في إحدى الفئتين

العلاقة بين مستوى الصعوبة ودرجة التمييز للفقرات

تعتمد قيم معاملات تميز الفقرات المتحصل عليها باستخدام طريقتي معامل

الارتباط المنصف للسلسلة الحقيقي، ومقارنة الأطراف العليا والدنيا على مستوى صعوبة الفقرات، فكلما ابتعدت قيمة معامل صعوبة الفقرة عن (0,50) قلت قيمة معامل التمييز للفقرة، وذلك لان الطريقتين السابقتين تقسم الدرجة الكلية إلى قسمين في الطريقة الأولى والى ثلاثة أقسام في الطريقة الثانية ثم يتم المقارنة بين القسمين الممثلين للفئة العليا والدنيا، مما يجعلها تتأثر بمعامل الصعوبة.

أما معاملات تمييز الفقرات المتحصل عليها باستخدام طريقة الارتباط المنصف للسلسلة فلا تتأثر بقيمة معامل الصعوبة حيث يتم الحصول على قيمتين مستقلتين لكل من معاملي الصعوبة والتمييز، وذلك لأن هذه الطريقة تعتبر الدرجات الكلية متغيرا متصلاً ولا تقسمه إلى قسمين مما يجعلها لا تتأثر بصعوبة الفقرة، وبالتالي تصبح هذه الطريقة المفضلة عند احتواء الاختبار على تدرج واسع من مستويات الصعوبة.

فعالية المشتتات (المموهات)

المشتتات عبارة عن الخيارات الأخرى غير الصحيحة للسؤال ذي الاختيار من متعدد، ويفترض في المشتتات التالي:

أ- إن تكون جذابة وبالذات للأفراد الذين لا يعرفون الاختيار الصحيح، أي يتم اختيار أي مشتت من قبل مفحوص أو أكثر بنسبة لا تقل عن 5٪ من المفحوصين.

ب- المشتتات عبارة عن إجابة خاطئة، لذلك فالمشتت الجيد هو الذي يكون عدد المفحوصين الذين يختارونه من الفئة العليا اقل من عددهم في الفئة الدنيا أي إنه (يميز باتجاه معاكس لتمييز الفقرة). وكلما كانت قيمة المموه بالسالب كان هذا دليل على أن المشتت جيد.

$$\text{معامل فعالية المشتت} = \frac{ن_ع - ن_د}{ن}$$

ن ع = عدد الأفراد في الفئة العليا الذين اختاروا المشتت
ن د = عدد الأفراد في الفئة الدنيا الذين اختاروا المشتت
ن = عدد الأفراد في إحدى الفئتين

إعداد الاختبار للاستخدام:

بعد أن ينتهي الباحث من جميع المراحل الأساسية لتصميم الاختبار يصبح من الضروري أن يضع الاختبار في صورته النهائية التي سيقدم بها إلى المفحوصين، من ذلك ما إذا كانت فقراته ستقدم في شكل بطاقات مستقلة تتضمن كل بطاقة فقرة واحدة أو ستوضع الفقرات في كتيب مستقل على أن يرفق بها صحيفة منفصلة للإجابة. كما يجب وضع تعليمات مناسبة للاختبار قبل تجربة الاختبار.

وفيما يلي بعض الأمور الواجب أتباعها عند وضع التعليمات للاختبار وتجربته

1- يجب شرح فكرة الاختبار شرحاً دقيقاً حتى لا يكون عدم فهم الأفراد لفكرة الاختبار سبباً في الإخفاق فيه.

2- يجب أن تبرز التعليمات الهامة وأن تكتب بخط واضح ومن المستحسن أن تكون حروفها أكبر من بقية التعليمات العادية.

3- من الأفضل أن تحتوي التعليمات على مثال أو نوع من التمرين للمختبر قبل كل اختبار.

4- يحسن أن تكون تعليمات الأجزاء المختلفة للاختبار، متسقة أو موحدة أن أمكن. فقد ثبت بالتجربة أن توحيد التعليمات يؤدي إلى دقة أكبر والى ألفة المختبر بالتعليمات وتمسكه بها بينما اختلافها قد يؤدي إلى بعض الأخطاء أو إلى التضحية بالدقة.

5- بالإضافة إلى أن التعليمات يجب أن تحتوي على طريقة الإجابة ومكانها وأسلوبها ويجب أن تحتوي أيضاً على الزمن المطلوب مراعاته عند الإجابة إذا كان الاختبار يتطلب ذلك.

6- من الأفضل إجراء تجربة استطلاعية على عينة صغيرة لغرض التعرف على مدى الاختبار ومدى فهم العينة لهذه التعليمات ويقصد بمدى ملائمة الاختبار أي مدى وضوح العبارات والمدة اللازمة للإجابة على الاختبار ومن ثم تعديل فقرات الاختبار في ضوء التجربة الاستطلاعية.

ومن الأمور المهمة التي ينبغي أن يراعيها واضع الاختبار هي وضع مفتاح للتصحيح إذ أن مفاتيح التصحيح تختلف باختلاف طريقة الإجابة أي قد تكون الإجابة على كراس الاختبار نفسه أو إذا كانت الإجابة على ورقة أجابه خاصة ودون كتابة شيء على كراس الاختبار. وطريقة الإجابة على الورقة أفضل بكثير وأكثر اقتصادية من الإجابة على كراس الاختبار نفسه إلا في الحالات الاستثنائية مثل اختبارات الدقة والتي تدخل فيها أرقام كثيرة أو رموز أو رسوم بيانية أو في حالة كون التلاميذ صغار العمر وخبرتهم محدودة في الاختبارات الموضوعية.

وأن مفاتيح التصحيح تكون بأنواع متعددة منها:

1- المفتاح ذو المروحة وهو الأكثر شيوعاً وهو أن يطبق المفتاح على حسب أقسامه الرأسية فيصبح على شكل مروحة ويلاحظ أن بالمفتاح قوائم بدرجات خام حولت إلى درجات معيارية.

2- المفتاح الشريطي ويشبه المفتاح ذو المروحة إلا إنه يختلف عنه في كون التصحيح الخاص لكل صحيفة من صحف الاختبار قد كتبت على حدة كما أن المفتاح قد كتب على ورق مقوى وليس على ورق عادي.

3- المفتاح المثقوب في مثل هذا المفتاح قد عملت ثقوب لبيان الكلمات أو الجمل الممثلة للإجابة في مفردات التكملة وهو يصلح لاختبارات التكملة.

4- المفتاح الشفاف وهو مصنوع من شفاف وشفافية تساعد على سرعة التصحيح وعلى تميز المفردات التي قد تصحح مرتين من جهة أخرى ويستخدم بكثرة في تصحيح الاختبارات الموضوعية.

5- المفتاح الكربون تصاحب هذا المفتاح ورقة الإجابة حيث تحدد أماكن الإجابات الصحيحة على ورقة مستقلة تلصق أطرافها في ظهر ورقة الإجابة حيث تكون مستقرة بالنسبة للمختبر ويطلى ظهر ورقة الإجابة بطلاء أسود بحيث يترك أثراً لأية كتابة أو علامة تسجل على ورقة الإجابة. بالرغم من سرعة هذا المفتاح في التصحيح إلا أنه كثير التكاليف.

6- المفتاح المطبوع فوق ورقة الإجابة وهو أن يلصق مفتاح التصحيح فوق ورقة الإجابة أي يطبع هذا المفتاح على ورقة الإجابة بعد أن ينتهي المختبرين من الإجابة.

7- المفتاح الآلي يمكن أن يستخدم المفتاح المثقوب في التصحيح بالآلات إذا أعدت أوراق الإجابة للتصحيح بالآلة وتكون وظيفة المفتاح الآلي هو تقسيم نقط الاتصال إلى دائرتين إحداهما للإجابة الصحيحة والأخر للإجابات الخاطئة كذلك قد يستخدم أحياناً مفتاح لاستبعاد المفردات حتى تطمس إجابات المفردات غير المرغوب فيها.

وبعد أن تمت كتابة المفردات وأعدت التعليمات يتم تجربة الاختبار تجربة مبدئية على عينة ممثلة للمجتمع المراد عمل الاختبار له وتسجل الملاحظات المختلفة فيما يتعلق بصلاحية التعليمات وفهم المختبرين لها وكذلك وقت الاختبار وتجمع كافة الملاحظات عن المفردات المختلفة وذلك تمهيداً لتعديل ما يحتاج منها إلى تعديل.

وأثر الانتهاء من التجارب الاستطلاعية الأولية يعمد المصمم إلى اختيار عينة أكبر من الأفراد المستهدفين في البحث بعد إجراء التعديلات المناسبة للاختبار ثم يجري مصمم الاختبار تجربة أساسية للاختبار Try-Out يطبق فيها الاختبار على عينة من الأفراد قوامها حوالي (400) فرد تحمل خصائص المجتمع المراد أعداد الاختيار له ومن خلال هذا التطبيق يتم بموجبه اختيار الفقرات الصالحة للصيغة النهائية للاختبار.

وعند ذلك يجب مراعاة نواحي متعددة عند اختيار الفقرات الخاصة بالصور النهائية للاختبار وهي:

1- صعوبة المفردات ومدى تدرجها.
2- قدرة المفردات على التمييز.
3- صدق المفردات.

وعن طريق إجراء التحليلات الإحصائية يتم تطبيق هذه الشروط وقد ذكرنا ذلك عند الحديث عن تصميم فقرات الاختبار في الفقرة الخاصة باختبار مستوى الصدق والصعوبة للفقرات.

تقنين الاختبار

يمكن القيام بعملية تقنين الاختبار بأتباع الخطوات الآتية:

أ- تحديد المجتمع الذي سيقنن عليه الاختبار تحديداً إجرائياً دقيقاً: حيث أن هذه الخطوة هي الأساس لضمان صحة الخطوات اللاحقة في عملية التقنين، وتتضمن هذه الخطوة تحيد أهم خصائص وسمات المجتمع الديمغرافية، من حيث الخصائص الجغرافية والسكانية والاقتصادية، وتوزيع الفئات العمرية فيه، ونوعية التعليم، والتركيبة الاجتماعية، حيث يتم من خلال هذه المعلومات تحديد العينة التي تمثل المجتمع تمثيلاً جيداً، كما أن هذه المعلومات تمثل خصائص المجتمع والتي على ضوئها يتم تعميم نتائج الاختبار على المجتمعات الأخرى.

ب- اختيار العينة الممثلة للمجتمع وتحديد أسلوب اختيارها: وتعتمد هذه الخطوة على الخطوة السابقة، حيث أن تحديد حجم العينة وأسلوب اختيارها يعتمد بشكل كبير على المعلومات المتوفرة عن مجتمع الدراسة كما تعتمد على الإمكانيات المادية والبشرية المتاحة، وبصفة عامة كلما كان حجم العينة كبيرا كلما كان أفضل وأقرب إلى التمثيل الجيد للمجتمع.

ج- التخطيط الجيد والمسبق لتطبيق الاختبار: وتتمثل هذه الخطوة في وضع خطة شاملة لتطبيق الاختبار تتضمن تحديد الإجراءات والخطوات التي سوف تتبع، وتجهيز جميع أدوات ومستلزمات الاختبار، مع وضع قوائم بأسماء الأماكن التي سوف يتم تطبيق الاختبار فيها (أسماء المدارس مثلا ") مع وضع برنامج زمني للتنفيذ.

د- تطبيق الاختبار: وتستلزم هذه الخطوة توحيد ظروف إجراء وتطبيق الاختبار لجميع أفراد العينة وذلك لتحقيق مبدأ تكافؤ الفرص أمام الجميع لضمان أن الفروق التي رصدها الاختبار تعود للفروق في أداء الأفراد فقط.

هـ- تحليل فقرات الاختبار: وذلك للتعرف على مدى فعالية فقرات الاختبار، ومدى إسهامها في الحصول على خصائص سيكومترية (الصدق والثبات) تتفق مع خصائص الاختبار الجيد، وتشمل هذه الخطوة التعرف على صعوبة كل فقرة من الفقرات، وقدرتها التمييزية، وتباينها، كما تشمل التعرف على مدى فعالية المشتتات لكل فقرة.

و- إيجاد الخصائص السيكومترية للاختبار (الصدق والثبات): وذلك للتأكد من أن الاختبار فعلا يقيس ما وضع لقياسه فقط، وان الفرق بين أداء الأفراد في الاختبار ناتج عن الأداء الحقيقي للأفراد وليس بسبب أخطاء القياس العشوائية.

ز- إيجاد معايير الأداء: وذلك من خلال أداء مجموعة التقنين وذلك بغرض توفير إطار مرجعي يمكن من خلاله الحكم على أداء الفرد في الاختبار. والشكل آلاتي يوضح الخطوات الأساسية لعملية تقنين الاختبارات النفسية:

الخطوات الأساسية لعملية تقنين الاختبارات النفسية

↓

تحديد الهدف والغرض من الاختبار

↓

هل يوجد اختبار معد مسبقاً يمكن أن يحقق الهدف

يوجد اختبار معد مسبقا يحقق الهدف ← الدراسة الحالية

لا يوجد اختبار معد مسبقاً يمكن أن يحقق الهدف

↓

تعريف السمة المقاسة تعريفاً إجرائياً محدداً

↓

تحليل المحتوى أو العمل أو المفهوم

↓

تحديد أوزان البنود

↓

صياغة بنود الاختبار

إخراج الصورة الأولية للاختبار من حيث

كراس الأسئلة – أوراق الإجابة – التعليمات – الملاحق

↓

تحديد مجتمع الدراسة

↓

تحديد عينة الدراسة وأسلوب اختيارها

↓

إجراء الدراسة الاستطلاعية للاختبار على عينة صغيرة من المجتمع للتعرف على مدى وضوح الأسئلة والتعليمات، والتعرف على الخصائص السيكومترية للاختبار إجراء التعديلات اللازمة للاختبار

↓

تطبيق الاختبار بصورته النهائية على عينة التقنين

↓

تصحيح الاختبار وتجهيز النتائج للتحليل الإحصائي

↓

تحليل بنود الاختبار وتشمل إيجاد صعوبة كل بند،
ومعامل التمييز وتباين البند وفعالية المشتت

↓

إيجاد الخصائص السيكومترية للاختبار (الصدق والثبات)

↓

إعداد معايير الأداء
إعداد دليل الاختبار وتجهيزه للنشر

الفصل الرابع

الصدق

يحتوي الفصل على:

- تعريف الصدق
- أهمية الصدق
- أنواع الصدق
- مقارنة بين أهم أنواع الصدق
- الطرق المختلفة لحساب الصدق
- تصحيح معامل صدق الاختبار
- العوامل المؤثرة في صدق الاختبار

الصدق:

عندما نريد اتخاذ قرار ما فإننا بحاجة إلى معلومات تساعدنا على اتخاذ هذا القرار، ولكن المشكلة التي تواجهنا هي كيف نختار الاختبار الذي سوف يحقق لنا غرضنا في اتخاذ هذا القرار وخاصة إذا عرض أمامنا عدد من الاختبارات تحت إطار الغرض نفسه، فيكون السؤال ما الاختبار الأفضل مما عرض؟ فإذا أردنا إدخال الطاولة الجديدة إلى غرفة في الطابق الثالث في المنزل وقبل أن نبذل جهداً

ونضيع وقتاً في حملها، نريد أن نتأكد من أنها ستدخل من باب الغرفة أم لا ولدينا ثلاثة مقاييس وهم المتر والكيلو جرام والمساحة فالأول للبعد والأطوال والثاني للأوزان والثالث للزمن فأي المقاييس سنتخذها لتساعدنا في تحقيق غرضنا؟ أن الإجابة بديهية وهي المتر.

وفي الجوانب الإنسانية إذا أردنا أن نعرف درجة فهم طلاب الصف الخامس الابتدائي لمعاني النصوص الأدبية في اللغة العربية فأننا سنصمم اختباراً ليقيس هذا الفهم فإذا طبقنا هذا الاختبار مرة أخرى على طلاب الصف السادس فلا شك أن نتائجه ستكون غير صادقة لأن الفئة التي طبقت عليها تختلف من حيث نضجها وخبرتها عن الفئة الذي وضع لها الاختبار، وإذا كانت نتائج الفئة الأولى التي وضع لها الاختبار عالية وقولنا بأن طلبة الصف الخامس لديهم قدرة عالية في اللغة العربية فستكون نتائجها خاطئة وغيرها صادقة لأن الاختبار لا يقيس قدرة الطلبة في اللغة العربية أنما يقيس فقط فهم معاني النصوص الأدبية.

ومن المقارنة بين مثال الطاولة السابق وتحديد الكم النسبي للصفات النفسية فلا تختلف الروائز من حيث الأهداف والغايات ولكن تعجز الروائز عن حمل الصفة النفسية لدى الفرد والأفراد وترتيبها ومقارنتها بينهم وذلك لأن عدد الأفراد كثيرون ولأن الصفة النفسية قدرة داخلية من الصعب ملاحظتها إلا من خلال نتائجها في السلوك وعلى هذا فعلى الرواز الاستعانة بأداة قياس صادقة ودقيقة تقيس ما يراد قياسه أي لا نقيس الأطوال بالكيلو جرام أو الأوزان بالمتر والقدرة اللغوية بمقياس الذكاء ونطبق اختباراً وضع للمعتوهين على الأذكياء. ولعل قد وضح لنا مفهوم الصدق ولكن ما تعريف الصدق؟

تعريف الصدق:

يعتبر الصدق من الخصائص المهمة التي يجب الاهتمام بها في بناء الاختبارات، فعندما يريد الباحث تصميم اختبار معين فلابد أن تكون هنالك ظاهرة سلوكية

معينة يقيسها الاختبار، كأن يقيس ظاهرة التكيف الاجتماعي أو الاتكالية أو القلق أو الذكاء أو التحصيل في موضوع دراسي معين. ويقوم الباحث بتحويل هذه الظاهرة السلوكية إلى عبارات يتألف منها الاختبار. وعندما يتأكد بطريقة علمية أن الاختبار يقيس الظاهرة التي يريد دراستها أو تشخيصها فعندئذ يعتبر الاختبار صادقاً. فالاختبار الصادق أذن هو ذلك الاختبار القادر على قياس السمة أو الظاهرة التي وضع لأجلها. فاختبار الاستعداد الدراسي يعتبر صادقا إذا كان قادرا على قياس الاستعداد الدراسي ويكون غير صادق إذا كان يقيس ظاهرة سلوكية أخرى.

وهناك تعريفات عديدة للصدق نستعرضها فيما يلي:

يعرف الصدق على إنه: (قياس الاختبار فعلاً وحقيقية ما وضع لقياسه) ويشير الصدق إلى (الدرجة التي يمكن فيها للاختبار أن يقدم معلومات ذات صلة بالقرار الذي سيبنى عليها).

وأنه (مقدرته على قياس ما وضع من أجله أو السمة المراد قياسها) أو (قدرته على قياس ما يدعى قياسه من جوانب سلوك الأفراد) أو (صلاحية الاختبار لقياس ما وضع لقياسه) ويعرف الصدق إحصائياً بأنه (نسبة التباين الحقيقي المرتبط أو المنسوب للسمة المقاسة إلى التباين الكلي ويشار إلى هذه النسبة بمعامل الصدق) ويورد صفوت فرج عدداً من التعاريف منها:

1- تعريف جيولكسون Gulikson: (الصدق هو ارتباط الاختبار ببعض المحكات محددا بذلك أن الارتباط بمحك خارجي في شكل معامل هو مؤشر الصدق). ولا يختلف هذا التعريف عن تعريف كلفورد أو جزء من تعريف الصدق لديه، إذ يذكر أن الصدق يوصف بتعبيرات الارتباط بين الاختبار وبعض مقاييس أو محكات الأداء في مواقف الحياة.

2- تعريف كورتن Cureton: (تقدير الارتباط بين الدرجات الخام للاختبار والحقيقة الثابتة ثباتاً تاماً). ويعتمد هذا التعريف على فكرة معامل الارتباط والتباين

المشترك بين الدرجات الخاصة بالاختبار وهذه الحقيقة، وهو يتطلب بالتالي تصحيح قياسنا للحقيقة الخارجية من عوامل عدم الثبات.

3- تعريف ليندوكوست Lindquist: (درجة الصحة التي يقيس بها الاختبار ما نريد قياسه). أو على إنه الدرجة التي تقترب بها بنجاح تام لقياس ما نريد قياسه، ورغم هذا التعريف يربط أيضاً بين الدرجة على الاختبار وبين هذا الذي نريد قياسه إلا إنه يتجنب هنا الإشارة إلى التعبيرات الإحصائية ومعاملات الارتباط.

4- تعريف أوجارتون Edgerton: (الصدق يشير إلى المدى الذي تكون به أداة القياس مفيدة لهدف معين). ويلغي بهذا الربط بين الدرجة ومحك آخر ويقتصر على محك الفائدة أو الاستخدام أو النفع من استخدام الدرجة.

5- تعريف كرونباخ Cronbach: (يقدر اكتمال درجة الاختبار للسمة المعينة والثقة من هذا التفسير بقدر صدق الاختبار). وهو يربط بذلك بين الدرجة على الاختبار وقدرتها التفسيرية.

6- يعرف كاتل الصدق بمعناه الواسع باعتباره قدرة الاختبار على التنبؤ ببعض الوظائف أو أشكال السلوك المحددة والمستقلة عن الاختبار والتي تعد محكاً لصدق الدرجة فيصنف أسلوب النظر إلى المحك من بعدين مستقلين:

أ- التجريد في مقابل العينات.

ب- الطبيعة في مقابل التخليق

أما Messick فيعرف صدق الاختبار بأنه تقييم شامل يوفر من خلاله الدليل المادي والمبرر النظري اللازمين لإثبات كفاية وملاءمة ومعنى أي تأويل أو فعل يبنى على درجة الاختبار.

وبعد هذه المرحلة الطويلة من التعريفات يتضح التكامل والتشابه بينهما فبعضها إجرائية وبعضها نظرية يجعل مؤشر الصدق هو الإتقان مع المحك وبعض التفسيرات والسمة ولكن هذه التعريفات تلخص في أن:

1- إن يكون الاختبار قادرا على قياس ما وضع لقياسه، أي أن يكون الاختبار ممثلا تمثيلا حقيقياً للقدرة التي صمم لقياسها.

2- إن يكون الاختبار قادرا على قياس ما وضع لقياسه فقط، أي أن الاختبار يقيس القدرة التي صمم لقياسها دون أن يخلط معها قياس قدرة أخرى غير مطلوب قياسها. فاختبار الذكاء مثلاً يجب أن يقيس الذكاء كل الذكاء ولا يقيس شيئاً آخر مع الذكاء.

وقد تختلف المصطلحات مثل صلاحية الاختبار أو صحته أو صدقه إلا أنها كلها تعني معنى واحد وأن كان المصطلح المشهور هو صدق الاختبار.

كما أن مفهوم الصدق مرتبط بنسبة التباين الحقيقي الخاص بالصفة التي نريد قياسها، فمن المعلوم أن التباين الكلي للدرجة الملاحظة يتكون من تباين الدرجة الحقيقية وتباين أخطاء القياس، ومفهوم الصدق مرتبط بالتباين الحقيقي، ويحدده نسبة التباين الحقيقي المناسب والناتج عن الصفة التي نريد قياسها، فقد يكون الاختبار رصد تباينا حقيقيا ولكن هذا التباين لا يعود جميعه للصفة التي نريد قياسها، فقد يكون الاختبار يقيس قدرة أخرى واستطاع أن يرصد التباين الحقيقي لها أيضاً، وبالتالي يصبح التباين الحقيقي للاختبار مكوناً من جزأين، جزء يعود إلى القدرة أو الصفة التي نريد قياسها وهو الذي يمثل صدق الاختبار، وجزء يعود إلى قدرة أو صفة أخرى استطاع أن يقيسها الاختبار حيث أنها اختلطت وتداخلت مع الصفة التي نريد قياسها.

وهذا يعني أن النتائج الخام لأي اختبار ينتج عنها تباين ينقسم إلى تباين حقيقي مناسب يعود للصفة التي نريد قياسها وهو الذي يمثل صدق الاختبار، بالإضافة إلى التباين الحقيقي غير المناسب الذي يعود إلى صفة أخرى تداخلت مع الصفة التي نريد قياسها وهو تباين مشوش لعملية القياس ويحد من موضوعيته، وهو نوع من التباين الثابت والذي يعود للأداة نفسها، ولا تتحكم فيه مؤشرات الثبات (يمثل أخطاء القياس الثابتة) بالإضافة إلى أخطاء القياس غير الثابتة والتي

تتحكم فيها مؤشرات الثبات. وبالتالي لكي يكون الاختبار صادقا فلابد أن يقيس القدرة التي صمم من لقياسها فقط ولا يقيس قدرات أخرى قد تتداخل معها.

3- أن يكون الاختبار قادرا على التمييز بين طرفي القدرة التي يقيسها، أي أن للاختبار القدرة التمييزية التي تسمح له بالتمييز بين الأداء المرتفع والأداء المتوسط أو الأداء المنخفض للأفراد.

أهمية الصدق:

والصدق له أهمية كبرى في تحديد قيمة الاختبار ومغزاه، فعنوان الاختبار قد لا يدل على ما يقيسه الاختبار فهو عادة ما يكون عبارة قصيرة تعرف بالاختبار بشكل عام، والصدق هو الذي يقرر صلاحية أي فرض علمي، أو في تحديد معنى مفهوم معين. وللصدق أهمية كبرى في الكشف عن المحتوى الداخلي للاختبارات النفسية، وفي الإفادة منه في الاختبار التعليمي المهني، وفي التنبؤ في حياة الأفراد التعليمية أو المهنية، توفيراً للجهد والمال والتدريب، فيطمئن الفرد بأنه يعمل في ميدان يتفق في ميدان مواهبه واستعداداته ومهاراته.

ونتيجة للاستخدام الضيق والمحدود لمفهوم الصدق من قبل الباحثين، ولأن الاختبارات والمقاييس تستخدم في مجملها لاتخاذ قرارات وإصدار أحكام متنوعة كلا منها يتطلب نوعاً معيناً من الدراسات للتحقق من صدقه، فقد أصدر علماء القياس التابعين للجمعية الأمريكية لعلم النفس (A P A) والجمعية الأمريكية للبحث التربوي (A E R A) كتيباً ضمنوه أهم معايير الاختبارات والمقاييس النفسية والتربوية وتصنيف للقرارات والأحكام المتنوعة التي يهدف إليها القياس بصفة عامة وهي:

1- تحديد كيفية أداء الفرد في الوقت الحاضر في نطاق شامل لمواقف سلوكية تمثلها مفردات المقياس تمثيلاً جيداً.

2- التنبؤ بالأداء المستقبلي للفرد أو تقدير مكانه في احد المتغيرات ذات الأهمية.

3- الاستدلال على درجة تملك الفرد سمة أو خاصية معينة من حيث هي تكوين فرضي يبدو أثرها في سلوك الفرد أو أدائه.

خصائص الصدق:

1- يرتكز الصدق على عاملين هامين هما الغرض من الاختبار الذي ينبغي أن يحققه، والفئة التي سيطبق عليها، فإذا كان الاختبار وضع ليقيس القدرة اللغوية يجب أن يقيس القدرة اللغوية قياساً دقيقاً شاملاً ولا يقيس شيئاً غير ذلك، وإذا كان الاختبار قد وضع لفئة معينة فيجب أن يطبق عليها وليس عليها وليس على غيرها. لأن كل فئة تختلف عن فئة أخرى من حيث النضج والخبرة وغير ذلك.

2- الصدق صفة نوعية أي خاص بما وضع لقياسه فإذا استخدمنا اختبار ليقيس الميل والاتجاه نحو اللغة فسيكون الصدق معتدل أو متوسط، وإذا استخدمناه للتنبؤ بنجاح الطلبة في الجغرافيا للعام القادم فإن النتائج ستكون ضعيفة.

3- الصدق صفة نسبية أو متدرجة: صفة الصدق ليست مطلقة ولهذا لا يصح القول أن الفحص يتصف بالصدق التام أو لا يتصف بالصدق أبداً ولكن تحدد درجة اتصافه أما بالارتفاع أو الاعتدال أو الانخفاض.

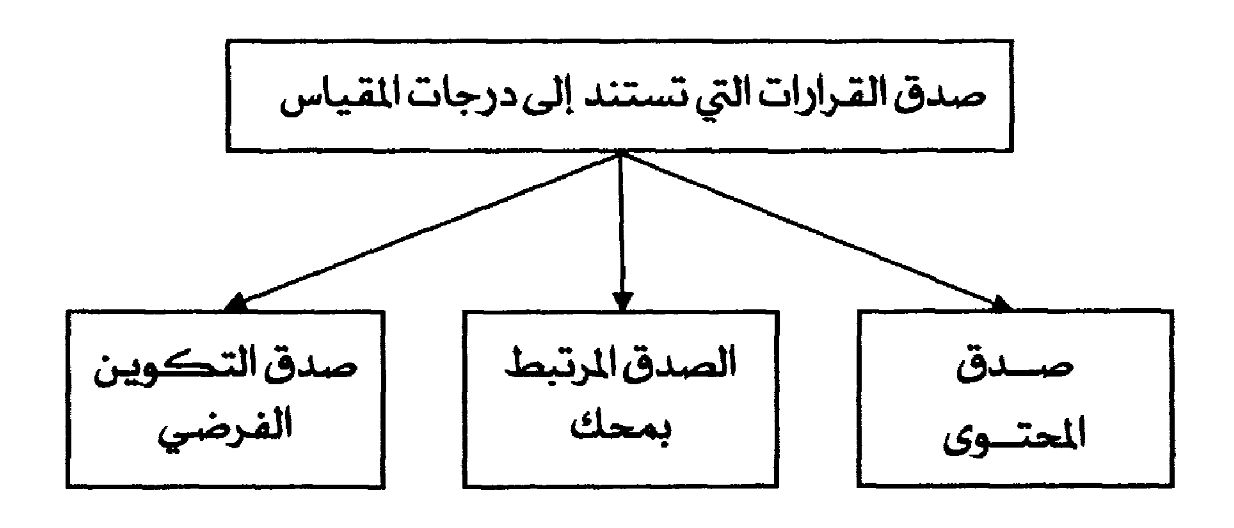
4- الصدق صفة تتعلق بنتائج البحث: فمن الشائع القول أن هذا الاختبار صادق والمعنى الأدق هو القول بأن نتائج الاختبار صادقة أي أن أحكامنا ستكون صادقة متى ما خلت النتائج من الخطأ.

5- يتوقف صدق الاختبار على ثباته: أي إنه إذا أعيد إجراء الاختبار في نفس الظروف فإنه سيحقق نفس النتائج، أما إذا اختلفت النتائج في كل مرة فهذا يعني أن الاختبار غير ثابت وغير صادق، أي أن الصدق يتوقف على ثبات

الاختبار والعكس غير صحيح أي لا يتوقف الثبات على الصدق فعلى سبيل المثال قد يكون مؤشر الميزان غير سليم فيعطي نفس الوزن لأوزان مختلفة إذاً فالنتيجة ثابتة ولكنها غير صادقة.

طرق التحقق من صدق الاختبار

عملية التحقق من صدق الاختبار لا تتم كما يحدث في التحقق من ثبات الاختبار عن طريق استخلاص مؤشر إحصائي أو معامل نسميه معامل الصدق، بل هو عملية يقوم فيها من صمم الاختبار بجمع الدلائل لتدعيم الاستنتاج الذي سوف يخرج به من درجات الاختبار، أي جمع دلائل تؤكد أن الاختبار الذي تم تصميمه يقيس ما صمم لقياسه فقط ولا يقيس شيئا آخر، أو إنه يقيس ما صمم لقياسه بالإضافة إلى شيء آخر، وعملية التحقق من الصدق تشمل عدة طرق يختلف استخدامها باختلاف نوعية الدلائل التي يرغب من صمم الاختبار الوصول إليها. وقد ارتبطت الطرق المختلفة للتحقق من صدق المقياس بأنواع القرارات والأحكام وهذه الطرق هي: صدق المحتوى والصدق المرتبط بمحك وصدق التكوين الفرضي وهذه الأنواع الثلاثة ليست مختلفة بل هي متكاملة وتعتبر أدلة على مدى صلاحية المقياس في المساعدة على اتخاذ القرارات وإطلاق الأحكام:



أولا: صدق المحتوى **Content Validity**:

يعرف صدق المحتوى بأنه (عمل أجرائي عقلاني أو منطقي يقوم على ضرب من المزاوجة بين المادة الدراسية وبنود الاختبار). ويعرف أيضاً بأنه الصدق الذي يتم عن طريق إجراء تحليل منطقي لمواد المقياس وفقراته وبنوده لتحديد مدى تمثيلها لموضوع القياس والمواقف التي نقيسها وبعبارة أخرى أن يقوم الفاحص بفحص مضمون الاختبار فحصاً دقيقاً

منظماً وتحديد ما إذا كان يشمل على عينة ممثلة لميدان السلوك الذي يقيسه، أي يكون الاختبار شاملاً لجميع أجزاء المحتوى ويمثله ويستطيع قياس مدى تحقيقه أهداف المحتوى، سواء أكان المحتوى مادة دراسية لصف معين وغير ذلك، عن طريق تحليل هذا المحتوى أو المقرر المدرس الذي ينوى قياسه وأعداد وسيلة لقياس الجوانب المختلفة لهذا العامل.

خطوات التحقق من صدق المحتوى:

يحدد الباحث الفاحص ما يريد قياسه تحديداً دقيقاً من موضوعات المحتوى والمادة الدراسية وأهدافها ويتعرف على المطلوب من التلاميذ من هذه المادة، ثم يحلل موضوع الاختبار تحليلاً شاملاً بين أقسامه ويرتبها حسب أهميتها ويحدد وزن كل قسم وكل ذلك تمهيداً لصياغة فقرات الاختبار بحيث يتفق عددها مع هذه الأوزان. فما تعرضه فقرات الاختبار هي التي تحدد حقيقة ما يقيسه الاختبار ولذا لابد من نظرة فاحصة إلى كل فقرة، فإذا وضع اختبار يقيس الكفاية في استخدام اللغة فكيف يمكن أن نتعرف إلى الدرجة التي يقيس بها هذا الاختبار التحصيل؟ والإجابة على ذلك نسأل أنفسنا أولاً ما المهارات والمعارف التي تؤلف الاستعمال السليم والفعال للغة والتي كانت أهدافاً لتدريس اللغة، بعد ذلك يتم فحص الاختبار وما يتطلبه من مهارة ومعرفة وفهم ثم لابد من إجراء المطابقة بين محتوى الاختبار وبين تحليل محتوى المادة، وأهداف تدريسها لترى بأية درجة تختلف نتائج هذا و محتوى الاختبار.

متى يكون صدق الاختبار عاليا:

إذا كان الاختبار شاملاً أي يقيس جميع ما درسه التلاميذ ولا يقيس بعضه فقط، كذلك إذا كان يقيس مدى تحقيق جميع الأهداف وليس بعضها، كذلك إذا كانت الأهداف معرفية ووجدانية ومهارية، ومن المعرفية التذكر والفهم والتقديم والتطبيق، يجب أن يشمل كل ذلك وقد تكون المطلوبة الحفظ كحفظ المحفوظات والأناشيد كما في الصفوف الأولى في المرحلة الأساسية، وقد تكون الأهداف الفهم والتفسير والنقد كما هو في تفسير القصائد في الصفوف الثانوية أي أن الاختبار يجب أن يمثل الأهداف والمحتوى بدقة. وإذا لم تتوفر هذه الدقة فسيضعف صدق الاختبار، فالاختبار العددي الذي يعتمد على الألفاظ أكثر مما يعتمد على الأعداد اختبار غير صادق والاختبار المكافئ الذي يعتمد على العمليات العددية أكثر من اعتماده على النواحي المكانية غير صادق من الناحية المنطقية ويكون الاختبار عالياً كلما كانت الموضوعات المراد قياسها محددة بدقة وتعتبر لائحة المواصفات مهمة في الاختبارات التحصيلية فهذه اللائحة تتضمن حصر الاختبار للموضوعات وتحديد أهمية كل منها فتتمثل في الاختبار مع ما يتناسب وهذه الأهمية.

أهميته وخصائصه واستخدامه:

يعتبر صدق المحتوى أهم أنواع الصدق عامة عندما يكون المجال محدوداً أو معروفاً فالذكاء على سبيل المثال أقل تحديداً في تحصيل الطالب في مادة اللغة العربية.

والصدق إجراء أولي بسيط يفيد في توليد الإحساس بالصدق ويعتبر أبسط معايير الصدق وأكثرها أولوية وتوليداً لليقين المباشر أو التلقائي بصدق في الأداة وقياسها لما تنوى لها أن تقيسه وهو عملية سهلة بإمكان أي معلم تطبيقها بهدف تصميم اختباراً لمادته وهذا يقود الباحث للإشارة والتنويه بأن وضع اختباراً قومياً دونما مراعاة خصائص الثبات المختلفة لم يتصف بالصدق، وكذلك لأن لكل معلم تصوراً خاصاً لمادته فالمدرس يدرس ويركز على مهارات تعتبرها هامة ويهمل

مهارات أخرى في رأيه غير هامة هذا ويستخدم هذا النوع عادة في دراسة صدق اختبارات التحصيلية.

حساب صدق المحتوى:

يقول البعض أن هذا النوع من الصدق لا يحسب له معامل صدق أنما نقارن الأسئلة والأهداف بالمحتوى الذي درسه الطلاب والأهداف التي وضعت لهذا المحتوى ويمكن إضفاء صفة الموضوعية أكثر عن طريق استشارة المتخصصين في المادة أو باستشارة الكتب ذات النظريات المتعارضة في تفسيرها للمادة غير أن هذه الموضوعية التي تصل إليها ليست كمية وليس أمام المصحح إلا اللجوء إلى معامل الاتفاق كما يلي:

قوام معامل الاتفاق:

1- جرد الباحثين من المقومات الأساسية التي ينوى تمثيلها في الاختبار.
2- محاولة كل باحث على انفراد صياغة تلك المادة في بنود مستقلة من عدد متفق عليه.
3- وضع جدول توزيع تكراري يمكن من تحديد درجة الصدق بالمعايير والمادة وذلك بتعداد النقط التي ضمنها الباحثون في الاختبار وحساب نسبتها إلى مجموع البنود.
4- تكرار العملية إلى أن يحصل الاتفاق وبدرجات عليا على توافق لا يقل عن 80٪ مثلاً.
5- توسعه شبكة الباحثين ومقارنة الواحد منهم بكل الباحثين المشتركين في عملية معامل الإتقان

وهذا يعبر عن توافق مصححي الاختبار فيما بينهم وبسبب ذلك التوافق لذلك يبقى الأمر ناقصاً. إلا أن فؤاد أبو حطب يورد لنا بعض الطرق للتحقق عن صدق المحتوى نلخصها فيما يلي:

1- إعداد صورتين متكافئتين من الاختبار تطبق إحداهما قبل الاختبار والأخرى بعده للتعرف على التحسن والتغير في الدرجات.

2- دراسة الأنماط والأخطاء الشائعة.

3- تحليل طرق العمل التي يستخدمها المفحوصون وذلك بإعطاء الاختبار فردياً مع توجيههم إلى التفكير بصوت مرتفع.

4- ويرى علام 2000 طرقاً أخرى تجريبية للتحقق من صدق المحتوى عن طريق تطبيق المقياس قبل بداية التعلم أو التدريب ثم تطبيق صورة مكافئة له بعد نهاية البرنامج للتحقق من حدوث أي تحسن في الدرجات.

ثانيا: الصدق الظاهري (السطحي) Face Validity

و(يعني البحث عما يبدو أن المقياس يقيسه) وهو المظهر العام للاختبار أو الصورة الخارجية له من حيث نوع المفردات وكيفية صياغتها ومدى وضوح هذه المفردات وكذلك يتناول تعليمات الاختبار ومدى دقتها ودرجة ما تتمتع به من موضوعية ودقة تحديد الزمن المناسب للاختبارات الموقوتة التي تعتمد على السرعة وعلى تحديد مستويات الصعوبة للاختبارات غير الموقوتة التي تعتمد على القوة وعلى نوع الأسئلة ومدى صلاحيتها للإشارة والاستجابة المناسبة من المختبرين. وهو يعني ما إذا كان الاختبار يبدو صادقاً في نظر المفحوصين من ناحية وفي نظر الفاحصين سواء أعد أعداداً فنياً أو لم يعد من ناحية أخرى. ويكون الفحص كذلك في رأي المفحوص إذا كان يقيس الغرض الذي من أجله قد تقدم للفحص أي إذا بدأ للمفحوص أن الأسئلة والأمثلة المستخدمة ذات علاقة بالوظيفة التي يريد قياسها.

ولتوضيح الفكرة فإن اختبار الصف الأول الابتدائي مثلاً يمكن أن يكون صادقاً ظاهرياً إذا كانت جميع فقراته ذات صلة بمادة الرياضيات للصف الأول الابتدائي ولا يقيس قدرته اللغوية مثلاً فإذا كانت صياغة الفقرات ذات طابع

لفظي فإن هذا يخفض من الصدق الظاهري ويثور الطلبة أحياناً في بعض الامتحانات لأن الفحص يبدو لهم ضعيف الصلة بما يمارسوه.

أهميته:

يجب على واضع الاختبار أن يراجع اختباره ليتأكد من توفر وجود شروط الصدق الظاهري. و الصدق الظاهري يكشف عن المفردات التافهة أو الضعيفة أو التي لا ترتبط كثيراً بالوظيفة المراد قياسها. وتأتي أهميته أيضاً بأنه يدخل في باب العلاقة الطيبة التي يكونها الفاحص مع المفحوص وهو من الأمور المرغوب بها فإذا بدأ الاختبار للفاحصين إنه لا معنى له أو لا علاقة له بالظاهرة المراد قياسها أو غير مناسب أو ساذج أو سخيفاً فيكون تعاون هؤلاء ضعيفاً، وأن كان الصدق الحقيقي للاختبار عالياً فينطوي هذا الصدق على سهولة الإمكانيات العملية لطبعه وتصحيحه وتفسير نتائجه.

تقييمه:

إن هذا النوع من الصدق يعد أقل أنواع الصدق جودة" حيث إنه من الملاحظ أن هذا النوع ليس إلا صدقاً ظاهرياً لا يلمس إلا سطح المقياس، فإن استخدام الباحث لهذا الصدق في اختبار صدق أداة معينة تقيس شيء ما وصممها وهي غير مقننة فإن الأساس المدعم لقبول المعلومات التي حصل عليها عن طريق هذه الأداة سيكون ضعيفاً. هذا ويعتبر البعض إنه لا يكاد يكون هذا النوع صدقاً على الإطلاق، وأن بدا في أعين الناس إنه صادقاً هذا ويمكن أن يتحسن الصدق الظاهري للاختبار عن طريق أعاده صياغة أسئلته بحيث تبدو أكثر ارتباطاً بموقف القياس الفعلي، ومع ذلك فإنه لا يمكن أن يكون بديلاً لأنواع الصدق الأخرى، وتحسينه لا يؤدي إلى تحسين الصدق الموضوعي.

حساب معامل صدقه:

ويتم التوصل إليه من خلال حكم المختص على درجة الاختبار للسمة المقاسة

وبما أن حكم المختص يتصف بالذاتية لهذا يعطى الاختبار للمحكمين، ويمكن تقييم درجته - الصدق الظاهري - من خلال التوافق بين تقديرات المحكمين. فعلى حسب توافق تقديرات المحكمين يكون مؤشر ضعف أو صدق الاختبار. وتقييمهم ما إذا كانت هذه الفقرات تتعلق بالشيء الذي نريد قياسه، ثم يقوم الواضع بعد ذلك بعمل تكرارات لاستجابات هذه المجموعة من المحكمين ويختار المقررات التي أتفق عليها أكثر عدد من المحكمين.

ثالثاً: صدق (التكوين الفرضي) المفهوم:

هذا النوع من الصدق مرتبط بالاختبارات النفسية بصفة عامة حيث ان معظم السمات في علم النفس افتراضية حيث لا يوجد اتفاق على تعريف أجرائي موحد لها، وصدق المفهوم يهتم بجميع الدلائل الضرورية والتي يتم بواسطتها التعرف على مدى قدرة درجة الاختبار في أن تكون مؤشرا على السمة التي يفترض أن يقيسها الاختبار، أي يتناول العلاقة بين نتائج الاختبارات والمقاييس وبين المفهوم النظري والذي يهدف الاختبار لقياسه.

وهذا النوع من الصدق يشكل المرحلة النظرية أو التمهيدية في تطوير الاختبارات والمقاييس، وهو موجه لخدمة الاختبار نفسه وذلك بمحاولة الانتقال من الشك في أن الاختبار يقيس السمة التي اعد لقياسها.

وكلمة (المفهوم) تدل على شيء قد يكون محسوس (كالشجرة) أو مجردة كصفة الذكاء والقدرة الموسيقية. وهو يتمثل في الارتباط بين الجوانب التي يقيسها المقياس وبين مفهوم هذه الجوانب ويسمى صدق المفهوم لأنه يقوم على تحديد المفاهيم أو البنى المقدمة للظاهرة المقاسة وبعبارة أخرى فهو نجاح الاختبار في قياس سمة أو قدرة معينة كالذكاء أو الانطواء أو الانبساط و يتوقف هذا النوع من الصدق على مقدار ما نحصل عليه من معلومات عن هذه السمة وخصائصها ومكوناتها ويسمى صدق المفهوم لأنه يقوم على تحديد المفاهيم أو البنى المقدمة

للظاهرة المقاسة واستنتاج المفاهيم والبنى التي تتولد عن منظومة البنى المقدمة من القدرات لدى الأفراد. ولتصمم مقياساً يقيس هذه الصفة أو القدرة يقوم الفاحص بتعريف الصفة التي يريد قياسها منطلقاً من إطار نظري، ثم يبدأ بتحليل المجال أو الناحية التي يريد قياسها تحليلاً يكشف عن عناصرها المختلفة الرئيسة ثم يحلل كل قسم ويجزئه إلى أجزائه، ثم يقدر النسب المئوية لكل جزء من كل قسم من هذه الأقسام، وهذا ما فعله الفرد بينيه عامي 1867م و 1911م وهو أول من وضع اختبار الذكاء وقد سار على الخطوات السابقة. وقد أورد كل من ساكس Sax وزيلر Zeller الخطوات التي يمكن من خلالها التوصل إلى دلائل تتعلق بصدق التكوين الفرضي (المفهوم) وهي:

1- تبرير أهمية التكوين الفرضي من الناحية التربوية والنفسية على أن يكون التكوين الفرضي معرفاً تعريفاً إجرائياً ويدل على سمة قابلة للقياس

2- الاستناد الكبير إلى نظرية تربوية أو سيكولوجية أو اقتراح نموذج منطقي يوضح المفاهيم والعلاقات القائمة بينها.

3- التمييز بين التكوين الفرضي والتكوينات الفرضية الأخرى المماثلة لها.

4- التوصل إلى أدلة من مصادر متعددة باستخدام الأساليب الارتباطية والتجريبية والمنطقية لتأكيد التكوين الافتراضي. وعندما تستخدم اختبارات متعددة في هذا الشأن فإنه يمكننا الحصول على الصدق التقاربي للتكوين الفرضي.

5- التوصل إلى أدلة نتأكد منها ان التكوين الفرضي لا يرتبط بعوامل وقتية أو دخيلة لكي نحصل على الصدق التمايزي للتكوين الفرضي.

6- إجراء تعديلات مستمرة في التكوين الفرضي بما يتفق والأدلة والمعلومات الجديدة المتجمعة.

أهميته:

وتتمثل قيمته في كونه يتناول بالدراسة أموراً مجردة لا تطالها الأدوات القياسية المباشرة كما هو الحال في قياس الشخصية والتذوق الأدبي والاتجاه العلمي والعادات، وقد يستخدم لقياس مظهر من مظاهر القدرة نفسها أو جانب منها، وقد تكون مهارة حسابية، وقد تكون قدرة رياضية، كما إنه يربط طريقاً بين القياس وبين التجريب في علم النفس، ويهتم علم النفس المعاصر كثيراً بهذا النوع من الصدق لأنه يمكنه من التخطيط التجريبي التحليلي لقياس الظواهر كان حرام قياسها مباشرة لأنها تسبب الأذى للإنسان أو بسبب عدم توفر أدوات القياس المباشرة. إن هذا النوع من الصدق يتخطى كل تلك الصعوبات والعقبات.

صعوبة تحديد صدق التكوين الفرضي (المفهوم):

إن محاولة قياس القدرات العقلية أو الاتجاهات أو القيم ليس بالأمر اليسير فهذه المفاهيم داخلية وليس لها شكل محسوس خارجي، ولا يلاحظها إلا الفرد أما الملاحظ الخارجي أو المقياس فلا يستطيع رؤية إلا أثارها الخارجية وحتى الفرد الذي يحمل هذه الصفة لا يحس إلا جانباً ضئيلاً منها.

ولو أن الفاحص أراد معرفة مدى دافع الأفراد لان يصبحوا أعضاء في مجتمعهم فهل إجاباتهم ستكون أجابه تمثل الحقيقة، قد يجهل الأفراد ما يريدون، وقد يكذبون لهذا تبرز صعوبة تحديد الصدق هنا.

الغرض من إيجاد صدق التكوين الفرضي (المفهوم):

والغرض من إيجاده هو معرفة طبيعة وقوة العناصر التي تؤثر في أداء المفحوص فعندما نريد أن نقيس الذكاء يجب أن يكون التعريف محدداً ودقيقاً بحيث نعزل العوامل الأخرى، فهل نريد قياس الذكاء بشكل عام أم السرعة في الإجابة أم التذكر أو غير ذلك.

محكات صدق التكوين الفرضي (المفهوم):

ويرجع ميخائيل أسعد صعوبة وضع محك لهذه المقاييس كمقياس الدافع مثلاً إلى سببين:

أولهما أن الدافع لا يظهر بثبات في موقف معين أو لحظة معينة فقد يظهر حيناً ويختفي حيناً آخر مما يصعب تصميم محك سلوكي يمكن الباحث من ملاحظة الناس في عدد كبير من المواقف وخلال أزمنة طويلة ومتعاقبة. **وثانيهما** وهو صعوبة أن يقوم على تحليل مبدئي واضح للمحتوى في اختبار دافع التحصيل مثلاً ليس شيئاً ملموساً حتى يقارن بشيء آخر أي بسلوك الفرد خلال فترات متعددة من زمن طويل.

ومع ذلك فتوجد محكات تكشف عن صدق المفهوم ومنها زيادة نسبة النجاح في الاختبار عند كبار السن عن صغار السن، فنحن نشك عندما ينجح التلاميذ في سن تسع سنوات بنسبة كبيرة في اختبار يقل فيه نجاح التلاميذ في سن عشر، وكذلك قدرة المقياس على بيان الفروق في الصفة المقاسة لدى فئات متضادة كالفرق بين الأفراد العاديين والمعتوهين، أو الفرق بين أعلى 10٪ في الصف وأدنى 10٪ من الصف نفسه.

وتوجد العديد من الأساليب أو المؤشرات الإحصائية وغير الإحصائية أو المنطقية المستخدمة في الكشف عن صدق التكوين الفرضي (المفهوم) منها:

أ- أساليب تعتمد على الارتباطات وتشمل:

1- دراسة العلاقة بين مجموعات مختلفة من الأفراد، وذلك للتعرف على قدرة الاختبار على التمييز بين مجموعتين عليا حصلت على أعلى الدرجات في الاختبار، ودنيا حصلت على أدنى الدرجات وذلك لكل فقرة من فقرات الاختبار.

2- إيجاد معاملات الارتباط بين درجات الاختبار الذي يفترض إنه يقيس تكوينا فرضيا معينا، ودرجات اختبار آخر ثبت بالعديد من الأدلة إنه يقيس التكوين الفرضي ذاته.

3- إجراء التحليل العاملي للتعرف على علاقة الاختبار باختبارات أخرى مشابهة وذلك للتعرف على الصدق التقاربي للاختبار، بالإضافة إلى علاقته باختبارات تختلف عنه وذلك للتوصل إلى الصدق التمايزي للاختبار.

4- إيجاد مؤشر الاتساق الداخلي للاختبار والمتمثل في معاملات الارتباط بين كل فقرة من فقرات الاختبار والدرجة الكلية.

ثانيا: أساليب تعتمد على التجريب

حيث تعتمد هذه الأساليب على التدخل التجريبي لإحداث تغييرات في درجات الأفراد في اختبار ما كوسيلة للتعرف على مدى تأثر الأداء بمعالجات أو متغيرات معينة، مما يساعد في تأكيد بعض التفسيرات المتعلقة بنتائج الاختبار أو رفضها.

ومن هذه الأساليب أسلوب تمايز الأعمار والذي يستخدم لمعرفة مدى تزايد درجات الاختبار بتزايد العمر وذلك من خلال تطبيق الاختبار على مجموعات عمريه مختلفة ثم التعرف على الفروق في الأداء بينها.

رابعا الصدق التجريبي أو الإحصائي Empirical Validity:

وهذا النوع من الصدق يعتمد طريقة مقارنة نتائج الاختبار الحالي أي الذي طبقة الفاحص مع نتائج اختبار آخر لنفس العينة وبنفس المواصفات ومن هذه المقارنة يوضح مدى صدق الاختبار الأول ويسمى الاختبار الأخر بالمحك وقد يكون فحص المدرس أو معدل الطالب الفصلي السنوي أو التراكمي أو معدل التخرج.

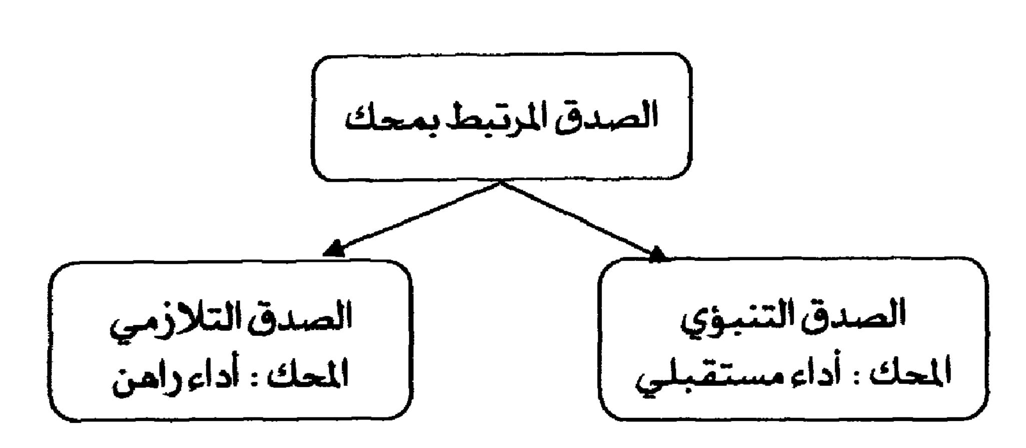
ويشترط في المحك اتصافه بـ (الموضوعية والصدق والثبات) وهناك طريقة أخرى وهي أن تتأكد من أن الاختبار قادر على التمييز بين الأفراد الذي طبق عليهم الاختبار وهم من ذوي القدرات المتفاوتة وذلك عن طريق تحليل بنود الاختبار وإزالة القنوات غير المميزة وذلك عن طريق حساب ارتباط درجة كل فقرة بالدرجة الكلية للاختبار.

وهناك طريقة أخرى كأن يطبق الاختبار على جماعتين متناقضتين كجماعة الأذكياء وجماعة الأغبياء فإن لم يميز بينهما فهو غير صادق وهو يعتبر من أهم أنواع الصدق بالنسبة للاختبار، ويقيس مدى قدرة الاختبار لقياس الوقائع الخارجية والتجريبية.

وهذا النوع من الصدق يحتوي على جميع أنواع الصدق التي تستخدم فيها الإحصاء والذي يحسب لها معامل ارتباط وهو يحتوي على الصدق التنبوئي والصدق التلازمي وسنتناول النوعين بشيء من التفصيل.

خامسا: صدق المحك

يرتبط هذا النوع بالأدوات التي نريد من خلالها تقدير مدى قدرتها على التنبؤ بأداء لاحق أو مستقبلي. وينقسم هذا النوع من الصدق إلى صدق تنبؤي وصدق تلازمي ويعتمد على معامل ارتباط بيرسون لتقدير معامل الصدق في كلا النوعين.



أ. الصدق التنبؤي Predictive validity:

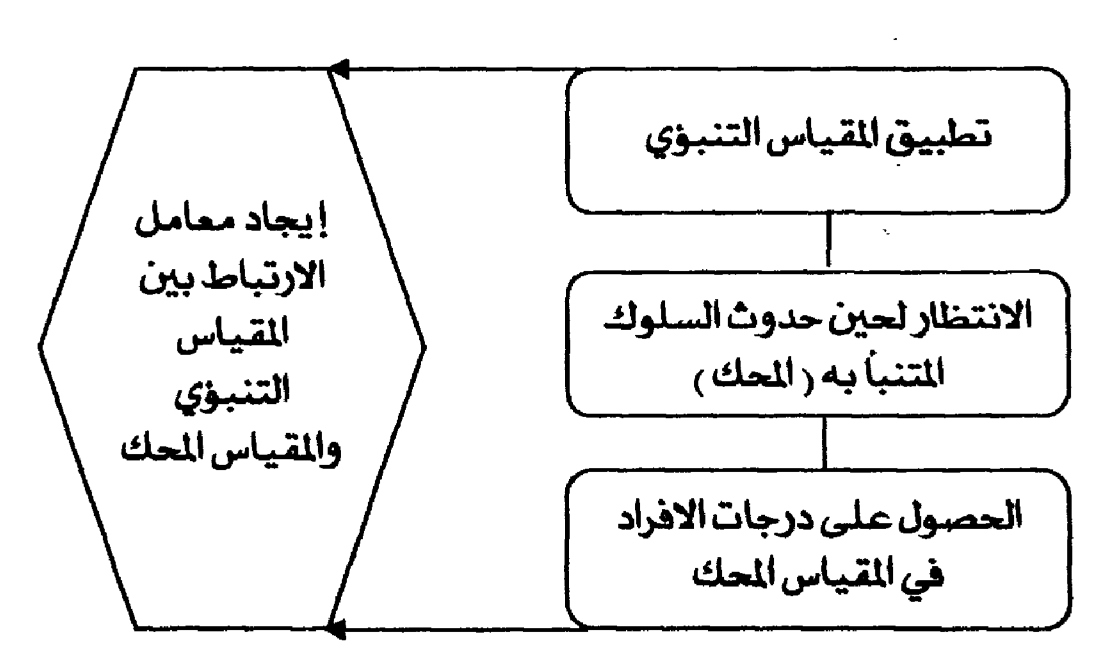
مفهومه:

ويقصد بهذا النوع من الصدق (قدرة الاختبار على التنبؤ بنتيجة معينة في المستقبل) ولمعرفة الصدق التنبؤي لاختبار موضوعي يجرى تجريب تطبيقيه على عينه من التلاميذ خلال العام الدراسي، ولكن الدرجات التي يحصل عليها أفراد

هذه العينة يحتفظ بها إلى أن يتم اختبارهم في وقت لاحق كمحك أو ميزان نقارن به أو نحدد المدى الذي تتفق فيه درجات الاختبار مع درجات المحك، وهذا الاتفاق يحدد لنا مدى تنبؤية الاختبار على أسس إحصائية.

حسابه:

وعلى العموم لابد لمحتوى هذا الاختبار الذي يمثل السلوك المقاس (المتنبىء) أن يكون ذا صلة بالسلوك المرتقب (المحك)، ولهذا فإن معرفة صدقه لا يعتمد على المحكمين، وهنا لابد من طرق إحصائية وهي معاملات الارتباط و يصلح هنا معامل ارتباط بيرسون للارتباطات حيث يحسب الارتباط بين الدرجات للاختبار الحالي بدرجات المحك أو الاختبار الأخر الذي يجتمع عنه المعلومات في فترة لاحقة لمعدل الفصل الأول أو معدل التخرج أو المعدل التراكمي. وعلى سبيل المثال اختبار القبول يكون المحك له معدل الفصل الأول في السنة الأولى ثم تحسب له معامل الارتباط بين نتيجة امتحان القبول ونتيجة اختبار الفصل الأول من السنة الأولى، ويكون المعامل عادةً أقل من الواحد الصحيح فإذا كان مثلاً 80٪ وجعلنا هذا الاختبار معياراً للقبول في الجامعة فأننا نتنبأ بنجاح 80٪ من الطلبة الذين هم فيه.



ب. الصدق التلازمي Concurrent Validity:

مفهومه:

وهو (نوع الصدق الذي يدل على وجود علاقة بين درجات الاختبار ومؤشر المحك التي نحصل عليها في نفس الوقت تقريباً) وهو يعتبر من أوسع أنواع الصدق انتشاراً والأكثر استخداماً في المرحلة الوقتية الراهنة وخاصة في علم النفس.

متى نستخدمه؟

ويستخدم عندما يريد المعلم استبدال اختبار جديد باختبار كان يستعمل وذلك عن طريق تطبيق الاختبار الجديد وحساب نتائجه وحساب معامل الارتباط لنتائج هذا الارتباط والاختبار الذي كان يستخدمه فإن كان الارتباط كبيراً أو 80٪ 0,8 عندها نقول أن هناك اتفاقاً بين الاختبارين ويمكنه استخدام إحداهما.

وهو يناسب الاختبارات التشخيصية التي تتعلق بمشكلات راهنة أكثر من التي تتعلق بمشكلات مستقبلية.

الفرق بين الصدق التلازمي والصدق التنبؤي:

إذا كان الصدق التلازمي يهتم بالإجابة على السؤال: هل هذا التلميذ متفوق في دروسه؟ فإن الصدق التنبؤي يجيب على السؤال: هل من المحتمل أن يصير هذا التلميذ متفوقاً؟ ولعل المثل السابق يوضح أن الفرق بين الصدق التلازمي والصدق التنبؤي هو فرق زمني فالمحك في الصدق التنبؤي يكون في زمن لاحق بينما في الصدق التلازمي هو في وقت مقارب ومصاحب للاختبار الذي وضعنا له هذا المحك.

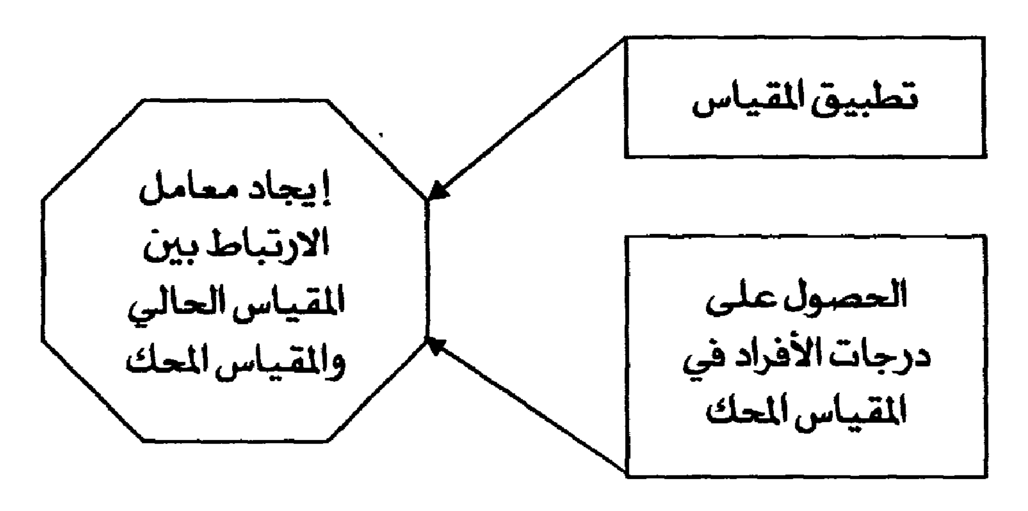
تقييمه:

وللصدق التلازمي مشكلات أهمها يتمثل في صعوبة تحديد علاقات السبب بالأثر، لأنه يدل على علاقة (اقتران) أكثر مما يدل على علاقة (علية) لهذا فإننا عندما

نستخدمه كبديل للصدق التنبؤي فإنه سوف يتضمن صعوبات منطقية ظاهرة، لأن التنبؤ يحتوي ويتضمن في جوهره هذه العلاقة السببية.

حسابه:

ويتم حساب هذا النوع من الصدق عن طريق معامل الارتباط بين نتائج الاختبار الذي بين أيدينا ونتاج المحك الذي قد يكون اختباراً أخراً على أن تجمع المعلومات عن نتائج المحك في الوقت نفسه أو في وقت متقارب لكليهما.



وعموما فإن قيم معاملات الصدق المرتبطة بمحك تتأثر بعدد من العوامل التي يجب أن نأخذها بعين الاعتبار عند تفسير الدرجات مثل:

1- تجانس العينة الذي يؤدي زيادته إلى انخفاض معامل الصدق وكلما كانت العينة غير متجانسة كلما زاد التباين وبالتالي زاد معامل الصدق، ذلك ان أحد المفاهيم الهامة لصدق المقياس هو قدرته على تمييز وإظهار الفروق الفردية لدى الأفراد في مجال سمة أو قدرة معينة. هذا من جهة ومن جهة أخرى فإن معامل الصدق هو في جوهره معامل ارتباط يتأثر كثيرا بمدى زيادة أو نقصان الفروق الفردية.

2- درجة ثبات المحك وأيضا ثبات المقياس حيث يجب أن يكون هذا المعامل عالياً لان القيمة الحقيقية لمعامل الصدق لا يمكن أن تتجاوز مؤشر الثبات (الجذر التربيعي لمعامل الثبات).

3- ذاتية وإطلاع المقيمين على درجات المقياس قد تؤثر في تقديراتهم للأفراد وبالتالي يتأثر معامل الصدق.

4- طول المقياس أيضاً له دور كبير في زيادة معامل الصدق فنظرا لان القيمة القصوى للصدق المرتبط بمحك تعتمد على مؤشر الثبات، فإن هذه القيمة تزداد بزيادة قيمة معامل الثبات، وزيادة قيمة معامل الثبات تعتمد على طول الاختبار، أي عدد فقراته، فكلما زاد هذا العدد زادت قيمة معامل الثبات إذا طبق الاختبار على مجموعة من الأفراد نفسها، غير أن هذا التأثير يكون أقل فيما يتعلق بقيم الصدق.

5- المدة الزمنية الفاصلة بين تطبيق المقياس التنبؤي ومقياس المحك لها تأثيرها، نظراً لأن معامل الصدق ينخفض بازدياد المدة الزمنية الفاصلة وذلك لتأثير الأخطاء العشوائية في درجات كل من المقياسين.

6- كما يفضل بل ينبغي أن نبدأ بعينة كبيرة من الأفراد تحسباً لفقدان بعضهم عند تطبيق مقياس المحك مما يؤدي إلى نقصان معامل الصدق.

سادسا: الصدق العاملي

(نبحث في هذه الطريقة لإيجاد الصدق عن عوامل مشتركة تقيسها عدة اختبارات صادقة لتحدد اشتراك هذه الاختبارات في قياس تلك العوامل)

فإذا أردنا معرفة مدى صدق اختبار يقيس صفة من الصفات فأننا نقوم بتحليل هذه الصفة إلى عواملها الجزئية، فالقدرة العقلية يمكننا أن نحللها إلى عوامل تتجزأ منها، القدرة اللغوية والقدرة المكانية والقدرة العددية فعندما نصمم اختباراً فأننا نصممه لكي يقيس جميع هذه العوامل الجزئية.

حسابه:

يعتمد على مدى اتفاق نتائج اختبارات أخرى تقيس هذه العوامل وهي اختبارات قد ثبت صدقها، فأنها حين ذلك نستطيع الحكم على هذا الاختبار الذي صممناه بأنه يتصف بالصدق العاملي.

تقييمه:

إلا أن هذه الطريقة لها عيوبها في تعدد معاملات الصدق العاملي للاختبار الواحد عندما تشبع بعوامل عديدة.

سابعا الصدق البنائي: Construct Validity

إن أي اختبار من الاختبارات يعتبر غير نقي إلى حد ما، ونادرا جدا ما يقيس بدقة الموضوع الذي يدل عليه عنوانه، لذلك فإن الاختبار لا يمكن تفسيره إلى أن نعرف العوامل التي حددت الدرجات عليه.

إن صدق البناء عبارة عن تحليل معنى .درجات الاختبار في ضوء المفاهيم السيكولوجية.

وهناك عدة طرق يمكن أن يلجأ إليها الباحث للوصول إلى هذا النوع من الصدق وأهمها:

1- إيجاد معاملات الارتباط مع اختبارات أخرى: ويتم ذلك عن طريق حساب معامل الارتباط بين الاختبار الذي نريد الوصول إلى صدق البناء فيه وبين الاختبارات التي تتشابه معه في الوظائف التي يقيسها والتي سبق التأكد من صدقها. فإذا افترضنا أن باحثاً أراد بناء مقياس للطموح بين الطلبة فإنه لكي يتوصل إلى صدقه يقوم بتطبيق اختبار آخر مستقل عنه يقيس المثابرة مثلا والتي تعتبر صفة من صفات الفرد الطموح وعندما يحصل الباحث على معامل ارتباط عال بين الاختبارين يتوصل إلى أن اختباره قد تحقق فيه صدق البناء.

2- إيجاد العلاقة بين درجات الفقرات والاختبار. وتفترض هذه الطريقة أن الدرجة الكلية للفرد تعتبر معيارا لصدق الاختبار، ويحاول الباحث عادة إيجاد العلاقة الارتباطية بين درجة كل فقرة من فقرات الاختبار والدرجة الكلية. وتحذف الفقرة عندما يكون معامل ارتباطها بالدرجة الكلية واطئاً، على اعتبار أن الفقرة لا تقيس الظاهرة التي يقيسها الاختبار بأكمله أو إننا قد نختار

مجموعتين متطرفتين من الأفراد بناءً على الدرجات الكلية التي حصلوا عليها في الاختبار، ثم نستخرج الفرق بين إجابات أفراد هاتين المجموعتين في كل عبارة من عباراته وتحذف العبارة التي لا تظهر تمييزاً واضحاً بين هاتين المجموعتين. وقد نقوم باستخراج العلاقة بين درجات الاختبارات الفرعية (إذا كان الاختبار يتألف من عدة اختبارات فرعية) وبين الدرجة الكلية للاختبار.

3- تقديرات الخبراء: وفي هذه الحالة نلجأ إلى الخبراء لتقدير الأفراد المبحوثين في الصفة السلوكية التي يقيسها الاختبار. فلو فرضنا أن أحد الباحثين صمم اختبارا لقياس التكيف الاجتماعي لدى طلبة المرحلة الإعدادية، وعرف الباحث التكيف الاجتماعي بأنه قدرة الفرد على ملائمة سلوكه مع ما تتطلبه المواقف الاجتماعية. فإنه لغرض أن يعرف فيما إذا كان الاختبار يقيس التكيف الاجتماعي أم لا، يختار مجموعة من الخبراء الذين يجهلون إجابة الطلبة على المقياس ويطلب منهم تقدير سلوك كل طالب في بعض المواقف ويجد العلاقة بين تقديرات الخبراء وإجابات الطلبة على المقياس فإن كانت عالية فإن ذلك يدل على إنه يتوفر في المقياس الصدق البنائي.

4- استخدام التحليل العاملي: ويهدف إلى معرفة مدى قياس الاختبار للظاهرة التي صمم الاختبار لأجلها. وتعتمد فكرته على حساب معاملات الارتباط بين كل عبارة والعبارات الأخرى، وتتجمع نتيجته المعاملات الارتباطية "مصفوفة ارتباطات" تنقسم إلى تجمعات، ويجمع بين كل مجموعة عامل أو أكثر، ونتيجة لهذه العملية فإن الاختبار يختزل إلى عدد صغير من العوامل أو السمات المشتركة والتي يطلق عليها "المكونات الأساسية" للظاهرة التي يقيسها الاختبار. ومن خصائص هذه الطريقة أنها تنقي الاختبار من العبارات الضعيفة التي يظهر أنها لا ترتبط بأي عامل أو مكون أساسي من مكونات الاختبار. وهذا يعني أن هذه العبارة لا تقيس الظاهرة التي وضع الاختبار لقياسها. والجدول التالي يوضح أوجه المقارنة بين أهم أنواع الصدق.

أوجه المقارنة بين أهم أنواع الصدق

جدول يوضح أوجه المقارنة بين أهم أنواع الصدق

النوع	المعنى	طريقة استخراجه	استعمالاته	أمثلة
صدق المحتوى	مدى قياس الفحص لما درسه التلميذ وما يتطلبه المنهج من أهداف مهارات.	قارن محتوى الفحص بمحتوى المنهاج والأهداف إلى أنواع السلوك والمهارات المتعلمة.	يحتاجه المدرس للتأكد من صدق اختبارات التحصيل.	فحص اختبار القدرة على الاختزال المعرفة هل يحتوي مضمونة بالفعل على عمليات الكتابة المكتبية.
صدق المفهوم	معرفة ماذا يقيس الفحص أو كيف نقيس أداء المفحوص حسب نظرية ما.	حدد العوامل التي تؤثر على نتيجة المفحوص بطريقة تجريبية ومنطقية وبالإمكان استخدام طرق أخرى من طرق الصدق.	يستعمل في بناء الاختبارات التي تقيس الصفات النفسية كالذكاء والقدرات.	فحص القدرة العددية
الصدق البنائي	كيف يمكن تفسير الدرجات على هذا الاختبار تفسيرا نفسيا؟	ضع فرضيات واختبرها تجريبيا بأي طريقة مناسبة.	استخدام الاختبارات للوصف أو البحث العلمي	يدرس اختبار الاستعداد الفني لتحديد المدى الذي تعتمد فيه الدرجات على التدريب الفني

النوع	المعنى	طريقة استخراجه	استعمالاته	أمثلة
الصدق التلازمي	مدى اتفاق نتائج الفحص مع نتائج فحص أخرى يحصل عليها من الفترة نفسها.	قارن بين نتائج الفحص مع نتائج فحص آخر يعطى في الوقت نفسه تقريباً.	كافة أنواع الاختبارات	مقارنة اختبار عقلي جماعي باختبار فردي
الصدق التنبؤي	استعمال نتيجة التلميذ في فحص ما للتنبؤ بأدائه أو نتيجته مستقبلاً.	قارن نتائج الاختبار مع نتائج تحصل عليها مستقبلاً.	اختبارات القدرات وكذلك اختبارات التحصيل المعيرة أو المقننة.	مقارنة اختبارات القبول بالكليات أو المعاهد بالدرجات التي يحصل عليها الطلاب فيما بعد

الطرق المختلفة لحساب الصدق:

لقد سبق الحديث عن كيفية حساب كل نوع من أنواع الصدق على حده وهنا نحاول أجمال تلك الطرق وإضافة طرق أخرى وهذه تتمثل فيما يلي:

1- طريقة معاملات الارتباط:

وهي من أكثر الطرق استخداماً وأدقها وتستخدم في الصدق التجريبي والصدق العاملي.

2- طريقة المقارنة الطرفية The comparison of extreme groups:

وفي هذه الطريقة يقسم الاختبار إلى قسمين أو ثلاثة، ثم يقارن بين متوسط القسم العلوي في الدرجات، ومتوسط القسم السفلي، وقد تكون مقارنة بين متوسط

درجات الأقوياء في الميزان مع متوسط درجة الضعفاء في نفس الاختبار بالنسبة لتوزيع درجات الاختبار، ثم نوجد الدلالة الإحصائية للفرق بين المتوسطات بواسطة النسبة الحرجة وتتصف هذه الطريقة بالسرعة والبساطة إلا أنها أقل الطرق دقة.

3- طريقة الجدول المرتقب Expectancy Chart:

وهي طريقة تقوم على أساس فكرة التكرار المزدوج وفيها يتم المقارنة بين التوزيع التكراري لدرجات الأفراد في الميزان بالتوزيع التكراري لدرجة الأفراد من الاختبار.

4- وقد ذكرت رمزية الغريب (1985) طرقا أخرى نلخصها فيما يلي:

أ- تعطى عدة اختبارات تقيس سمات واحدة لعينة من الأفراد ثم تحسب ارتباط كل اختبار منها بالاختبارات الأخرى ومن ثم نتعرف على الاختبارات الأكثر ارتباطاً فهي الأكثر صدقاً.

ب- طريقة الاعتماد على المحكمين وهم من المتخصصين في مجال القياس الاختبار.

ج- طريقة الصدق الذاتي: وهذه الطريقة تعطينا معامل الارتباط بين الدرجات التجريبية والدرجات الحقيقية.

العوامل التي تؤثر في صدق الاختبار:

أولا: عوامل تتعلق بالتلميذ:

1- قلق التلميذ أثناء أدائه الامتحان تؤثر على أجابته وفي نتيجة الاختبار وصدقه.

2- قد يمتحن التلميذ ويغش فلا تعبر النتائج عن مستوى التلميذ وقدرته.

ثانيا: عوامل متعلقة بإدارة الاختبار:

1- هناك عوامل بيئية كالحرارة المرتفعة والبرودة الشديدة والضوضاء تؤثر في الإجابة.

2- أن سوء الطباعة والأخطاء المطبعية وكيفية صياغة الأسئلة كلها تؤثر في الإجابة.
3- التعليمات غير الواضحة وخاصة في الأسئلة الموضوعية الحديثة التي تتطلب طرقاً خاصة في الإجابة كالاختيار من متعدد قد لا يفهم الممتحن كيف يجيب فتؤثر على نتيجة اختباره.
4- استخدام الاختبار في غير ما وضع له، كأن يوضع الفاحص اختباراً في النحو ويريد يعرف مستوى التلاميذ اللغوي.
5- تطبيق الاختبار على مجموعة لم يوضع لها الاختبار، كأن يطبق الاختبار على المتفوقين وهو قد وضع للمستويات العادية المتوسطة.

ثالثا: عوامل متعلقة في الاختبار نفسه:

1- صعوبة الأسئلة وسهولتها أي أن تكون الأسئلة فوق وأعلى مستواهم أو أدنى.
2- غموض الأسئلة، فقد تفسر تفسيرات أخرى مختلفة.
3- صياغة الأسئلة. كأن تكون الأسئلة موحية بالإجابة.
4- طول الاختبار.
5- ثبات الاختبار.
6- ثبات الميزان.
7- اقتران ثبات الاختبار بثبات الميزان.
8- التباين.
9- صدق مقياس المحك.
10- مدى تمثيل الاختبار باعتباره عينه من المثيرات للسلوك المطلوب قياسه.
11- طريقة حساب معامل الصدق.
12- تجانس عينة التقنين.

العلاقة بين الثبات والصدق:

قبل أن نبحث في العلاقة بين الصدق والثبات يجب أن ندرك بعض أوجه الشبه

والاختلاف فيما بينهما، فالصدق والثبات يعتبران أهم سمتين للمقياس الجيد حيث ان الثبات يبحث في مدى اتساق وثبات مفردات هذا المقياس بينما الصدق يتعلق بالهدف والغرض الذي بني من أجله المقياس. من ناحية أخرى فإن ثبات المقياس يتأثر بالأخطاء العشوائية غير المنتظمة.

بينما صدقه يتأثر بجميع أنواع الأخطاء سواء كانت عشوائية أو منتظمة، ومن جانب ثالث فإن الصدق سمة لا تتعلق بالمقياس نفسه بقدر تعلقها بتفسير الدرجة المستخرجة منه، وهذا يعني مدى فائدة أداة القياس في اتخاذ قرارات تتعلق بغرض أو أغراض معينة.

ولأن المقياس لا يمكن أن يكون صادقاً إذا لم تتسم مفرداته بالاتساق والثبات (حيث أن قيمة مؤشر الثبات تعد الحد الأعلى لقيمة معامل الصدق) فإن تفسير الدرجة المستخرجة من مقياس غير ثابت سيكون بالتأكيد تفسيراً خاطئاً.

كما أن لقيم معامل الثبات تأثيرها في درجة العلاقة بين المقياس التنبؤي والمحك كما سبق وان أوضحنا، ثم أن قيم معامل الاستقرار تسهم في زيادة الثقة بصدق المفهوم الذي يهتم في أحد جوانبه بالسمات المميزة بالثبات النسبي. ويعتبر معامل التجانس أو الاتساق الداخلي متى ما وجد في مقياس ما دليلاً على صدق المفهوم لهذا المقياس.

وأخيراً لعلنا نلاحظ بأن المقياس الصادق لابد وبالضرورة أن يكون ثابتاً بينما قد نجد مقياساً ثابتاً يتميز باتساق مفرداته ورغم هذا يكون غير صادق، حيث ان الصدق يتصل بالغرض والهدف من المقياس فمثلاً اختبار ذكاء الأطفال بالرغم من إنه يتمتع بدرجة عالية من الثبات إلا إنه في حال استخدامه لقياس ذكاء الكبار فلن يكون صادقاً في النتائج المستخرجة ولا يستطيع أن يعطي دلالات وتفسيرات صحيحة عن ذكاء الكبار.

الفصل الخامس

الثبـــات

يحتوي الفصل على:

- تعريف الثبات
- أنواع معامل الثبات
- نظرية الثبات
- العوامل المؤثرة في ثبات الاختبار
- أساليب حساب معامل الثبات

الثبات Reliability :

نشأ مفهوم الثبات نتيجة لظهور أعداد كبيرة من الأدوات والاختبارات التي تقيس مختلف نواحي السلوك الإنساني وذلك بعد الحرب العالمية الثانية، وبالتالي نشوء شركات القياس التي أخذت تتبنى هذه الاختبارات وتقوم بتسويقها، والتي واجهت مشكلة تتمثل في الإعداد الكبيرة من الاختبارات حيث أصبح لكل مفهوم أو مصطلح نفسي عدد كبير من الاختبارات التي تقيسه، والسؤال الذي طرح نفسه هو أي من هذه الاختبارات أجود وأفضل من حيث دقته في قياس الظاهرة التي يراد قياسها، وهذا التساؤل جعل شركات القياس تتجه إلى إيجاد المعايير الأساسية

التي يحتكم إليها لتحديد جودة الاختبار، وكان من أهم المعايير والتقنيات مفهوم الثبات، حيث بدأ علماء القياس في حسابه بطرق عديدة تقوم الفكرة الأساسية لمعظمها على معامل الارتباط.

ويعد مفهوم الثبات من المفاهيم الجوهرية في القياس، ويمثل مع مفهوم الصدق من الأسس المهمة للاختبارات والمقاييس، إلا أن مفهوم الصدق أشمل وأهم من مفهوم الثبات، إذ يمكن القول أن كل اختبار صادق هو بالضرورة ثابت، لكن ليس كل اختبار ثابت هو صادق بالضرورة. حيث أن الاختبار الصادق الذي يقيس فعلاً ما أعد لقياسه، فإن درجته ستكون معبرة عن الأداء الحقيقي أو القدرة الفعلية للفرد، لذلك مادامت الدرجة على المقياس الصادق تعبر عن هذه الوظيفة بدقة، فأنها تكون ثابتة في الوقت نفسه أي متسقة داخلياً في تعبيرها عن مقدار الوظيفة، ومستقرة عبر الزمن في تعبيرها وتقديرها لهذه الوظيفة إلا أن هذا لا يعني إنه يمكن الاستغناء عن الثبات في أعداد وبناء الاختبارات والمقاييس، وذلك لعدم توفر مقاييس واختبارات تتسم بالصدق التام، بالإضافة إلى أن تقدير الصدق يخضع عادة لاعتبارات عملية متعددة تختلف من حالة لأخرى ومن مجتمع لأخر مما ينبغي تقدير الثبات أضافه لتقدير الصدق قبل التطبيق النهائي للاختبار أو المقياس على عينة معينة أو مجتمع معين.

ويوفر معامل الثبات Reliability coefficient الكثير من المؤشرات الإحصائية للصفة أو الظاهرة المدروسة، والتي من خلالها يمكن الحكم على دقة المقياس الذي أستخدم في القياس، بالإضافة إلى أن تقدير ثبات الاختبار يزود الباحث بمعلومات أساسية للحكم على نوعية تكنيك الاختبار ومدى صلاحيته ودقته واتساقه فيما يزودنا به من بيانات عن الصفة أو الظاهرة المدروسة.

وتبرز أهمية الثبات في كوننا لا نملك مقاييس واختبارات سيكولوجية وتربوية مطلقة الأحكام والدقة، كما لا نمتلك التحكم التام والضبط الدقيق لمواقف القياس، مما تتسرب من خلالها بعض المتغيرات الدخيلة المؤثرة في القياس، وبالتالي فإن

الدرجة التي نحصل عليها من المقياس أو الاختبار لا تكون معبرة بدقة عن الظاهرة التي نقيسها، سواء أكانت سمة أو قدرة أو غيرهما، إذ تتضمن الدرجة التي نحصل عليها دائماً قدراً من الخطأ سواء أكان خطأ موجباً في شكل زيادة في الدرجة عما يستحق الفرد نتيجة لقدرته الحقيقية، أم خطأ سالباً في شكل نقصان في الدرجة، مما يكون تقييم أداء الفرد أقل من الواقع الذي يستحقه.

وإذا حصلنا على درجات دقيقة وصحيحة بواسطة أي اختبار أو مقياس. بدون أخطاء يمكن أن نرجعها إلى طبيعة ونوعية فقرات الاختبار، أو إلى تخمينات وتحيزات الفرد عند الإجابة على الاختبار، أو إلى الأخطاء التي تحدث بسبب سوء الفهم من قبل الشخص المختبر، أو إلى العوامل الأخرى التي يمكن أن تؤثر على أدائه الحقيقي، فعند ذلك سيمتلك الاختبار ثباتاً تاماً (كامل) مما سيكون معامل ثبات مثل هذا الاختبار يساوي (1).

بيد أن في الواقع، لا يوجد اختبار تحصيلي أو اختبار عقلي أو نفسي أو أي قياس، يمكن أن يحصل على درجة ثبات تامة (1)، لأنه لا يمكن تجنب جميع هذه الأخطاء، لذلك فإن مهمة الباحث أو المختص هو تقليل الأخطاء المحتومة في القياس إلى أدنى درجة معقولة. وعليه فإن الدرجة التي نحصل عليها من أي اختبار لا تعبر عن الأداء الحقيقي للفرد فقط، بل هي تمثل الأداء الحقيقي للفرد والشوائب الدخيلة في الدرجة سواء كانت موجبة أو سالبة، بمعنى آخر أن درجة الفرد على الاختبار أو المقياس تعبر عن التباين الحقيقي للفرد True variance وتباين الخطأ Error variance. ولذلك يؤكد (جيلفورد) guilford على ضرورة حساب ثبات الاختبار كي نحدد الدرجة الحقيقية أو التباين الحقيقي للاختبار، لأن معامل الثبات يوضح نسبة التباين الحقيقي في الدرجة المحسوبة على الاختبار وعليه فإن مهمة أساليب حساب الثبات المختلفة هي أن توفر تقديراً جيداً لحجم التباين الحقيقي في الدرجة الكلية المستخرجة مع الإشارة إلى تباين الخطأ.

تعريف الثبات:

يقصد بثبات الاختبار أن تكون أدوات القياس على درجة عالية من الدقة والإتقان والاتساق والاطراد فيما تزودنا به من بيانات عن السلوك المفحوص. ويرى علام 2000 إنه متى ما كانت أداة القياس خالية من الأخطاء العشوائية، وكانت قادرة على قياس المقدار الحقيقي للسمة أو الخاصية المراد قياسها قياسا متسقا وفي ظروف مختلفة ومتباينة كان المقياس عندئذ مقياساً ثابتاً. ولهذا فإن الثبات هو الاتساق والدقة في القياس. ويمكن أن نوضح مفهوم الثبات من خلال مقابلته مع مفهوم الصدق حيث أن مفهوم الثبات يعني الاتساق في مجموعة درجات الاختبار التي فعلاً قاست ما يجب قياسه.

أما مفهوم الصدق فيعني الدقة في درجات الاختبار التي يفترض أن تقيس ما يجب قياسه.

أما التعريف الإجرائي لمعامل ثبات مجموعة من درجات مجموعة من الممتحنين فهو معامل ارتباط بين تلك المجموعة من الدرجات ومجموعة أخرى من الدرجات على اختبار مكافئ، تم الحصول عليها بشكل مستقل من أفراد نفس المجموعة وهذا التعريف الإجرائي يتطلب قياسين أو أكثر لنفس الصفة أو الظاهرة لكل عضو من أعضاء المجموعة بواسطة اختبارات متكافئة ويستخدم مفهوم الثبات بمعناه العام ليدل على مدى اعتماد الفروق الفردية في درجات الاختبار على أخطاء الصدفة المتضمنة في القياس من ناحية، وعلى الاختلافات الحقيقية في الصفة أو الخاصية موضع الدراسة من ناحية أخرى. والثبات بهذا المعنى يشير إلى نسبة التباين الحقيقي في الدرجة التي يحصل عليها الفرد من اختبار ما، ولذلك فإن معامل الثبات هو هذه النسبة من تباين الدرجة على المقياس التي تشير إلى الأداء الفعلي للمفحوص، مما يتضمن هذا المعنى تضيف الدرجة التي يحصل عليها الفرد إلى مكونين رئيسين هما التباين الحقيقي لأداء الفرد على المقياس وتباين الخطأ أو الدرجة الزائفة نتيجة لشوائب القياس.

وهناك بعض المؤشرات والملاحظات التي ينبغي ذكرها كي تكون الصورة واضحة عن معنى وأهمية تقدير الثبات أو الاختبار، ومنها:

أن الثبات ضروري، لكنه ليس هو الحالة الكافية لتحديد نوعية ودقة الاختبار، حيث قد نحصل على درجات ثبات عالية، لكننا قد نقيس بدقة شيئاً ما لا علاقة له بالموضوع أو الظاهرة المراد قياسها. أما إذا حصلنا على درجات غير ثابتة (Unreliable) فإن المقياس لا ينجح في قياس أي شيء تماماً من الظاهرة أو الصفة المدروسة. مما ينبغي أن نتأكد أولاً من موضوعية وصدق المقياس في قياس السمة أو الظاهرة المراد قياسها ومن ثم التعرف على درجة ثباته.

أن الاختبار أو المقياس نفسه بوصفه مجموعة من الفقرات أو الأسئلة لا ثبات له، بينما الثبات بالأحرى إلى الاختبار عندما يطبق على مجموعة أو عينة من الأفراد ولذلك فإن الثبات هو ثبات الإجابة على الاختبار، أي ثبات أداء الفرد على الاختبار، ولهذا فإن ما يحسب ثباته هو عينة استجابات مجموعة من الأفراد.

إذا كانت الفروق بين درجات الفرد نفسه ذات صلة قليلة بالفروق بين درجات الأفراد المختلفين فإن الاختبار يتجه إلى إعطاء ثبات عالي، في حين إذا كان الفرق بين درجات الفرد نفسه ذات صلة كبيرة بالفروق بين الأفراد فإن الدرجات تظهر ثباتاً واطئاً.

بما أن الثبات يعني ثبات عينة الاستجابات من مجموعة الأفراد، وبما إنه لكل مجموعة من الأفراد خصائصها، فمن الضروري أن نحصل على أكثر من معامل ثبات للاختبار الواحد لدى أكثر من عينة.

إن حصولنا على معامل ثبات لأداء عينة من الراشدين على اختبار معين، لا يعني أن ثبات أداء المراهقين أو الأطفال سيكون في نفس المستوى من الثبات، كما لا يجوز مطلقاً أن نقبل ثبات أداء عينة من الذكور على إنه مطابق لثبات أداء عينة من الإناث على نفس الاختبار. كما تلعب الفروق الحضارية والتعليمية وغيرها

دوراً هاماً في التأثير على ثبات المقاييس أو الاختبارات. لذلك ينبغي أن نقدر معامل ثبات أي اختبار قبل تطبيقه على عينة الظاهرة المدروسة مهما كان ثابتاً على عينات أخرى أو في بيئة أخرى أو في وقت سابق.

تتعدد أساليب حساب معامل الثبات ويختص كل أسلوب أو طريقة منها لتقدير نوعية معينة من تباين الخطأ، وبالتالي نوعية معامل الثبات المراد تقديره. وعلى الباحث أن يحدد نوع معامل الثبات قبل استخدام الأسلوب الذي يقيس ذلك النوع من الثبات والجدول التالي يوضح أنواع معامل الثبات وتباين الخطأ وطريقة تقدير كل نوع منها.

	نوع معامل الثبات	طريقة التقدير	تباين الخطأ
1	معامل الاستقرار	أعادة نفس الاختبار على العينة نفسها بعد فترة زمنية	التغيرات المؤقتة
2	معامل التكافؤ	تطبيق الاختبار بصورتيه المتكافئتين في نفس الوقت	عينة محتوى الاختبار
3	معامل الاستقرار والتكافؤ	تطبيق الاختبار ثم تطبيق الصورة المكافئة له بعد فترة مناسبة	التغيرات المؤقتة وعينة المحتوى
4	معامل الاتساق الداخلي	التجزئة النصفية للاختبار، ثم يصحح بمعادلة (سبيرمان – براون) أو غيرها.	عدم تجانس نصفي الاختبار
5	معامل الاتساق داخل الأسئلة	تحليل الفقرات باستخدام أحد معادلتي (كيودر - ريتشارد سون) أو غيرهما.	التجانس الكلي لفقرات الاختبار

نظرية الثبات:

تتولى نظرية الثبات مهمة تحليل طبيعة كل من التباين الحقيقي وتباين الخطأ والعلاقة بينهما لتوفر المنطق الأساسي الذي تقوم عليه الطرق المختلفة لحساب ثبات الاختبار، حيث إن:

الدرجة الكلية المستخلصة من المقياس=

التباين الحقيقي لأداء الفرد على المقياس + تباين الخطأ نتيجة لشوائب المقياس

.... معادلة (1)

أما إذا استخدمنا مقياس محكم تماماً في اختبار فرد ما تحت ظروف مثالية للقياس فسنجد أن التباين الحقيقي هو عبارة عن الدرجة التي يمكن أن يحصل عليها هذا الفرد، غير أن هذا المستوى من القياس وهذه الظروف لا توجد في الواقع العملي مما لا يمكن أن تكون الدرجة الكلية ممثلة للتباين الحقيقي لدرجة الفرد. والطريقة الأخرى التي يمكن أن تلجأ إليها لتحديد الدرجة الصحيحة للفرد (أي التباين الحقيقي) هي أن نختبر نفس الفرد عدداً كبيراً من المرات بنفس المقياس وتحت نفس الظروف أو ظروف مشابهة ثم نحسب متوسط الدرجة التي يحصل عليها الفرد في هذه الاختبارات المتعددة وبهذا تكون الدرجة الحقيقية أو التباين الحقيقي عبارة عن متوسط درجات القياسات المستقلة بواسطة نفس الاختبار.

والمنطق الذي يؤيد اختبار متوسط الدرجة في هذه الاختبارات المختلفة هو التباين الحقيقي هو أن تباين الخطأ أو الجزء الزائف من درجة الفرد على المقياس قد يكون سالباً أو موجباً، ووفقاً لخصائص المنحني الاعتدالي فإن متوسط هذه الدرجات أو المقادير الخاطئة في عدد كبير من مرات الاختبار سيكون صفراً. وبذلك يعبر متوسط أداء الفرد في هذا العدد الكبير من الاختبارات عن درجته الحقيقية لأن الدرجات الزائدة الموجبة والدرجات الناقصة السلبية ستتعادل وتصبح صفراً وتبقى فقط الدرجة المعبرة عن أدائه الفعلي أي المتوسط. إن هذا المنطق سليم نظرياً من ناحيتين:

الأولى: أن المجموع الجبري لتباين الخطأ سيكون صفراً في عدد كبير للغاية من مرات الاختبار للفرد الواحد وأن كان من المحتمل أن لا يكون هذا المجموع صفراً في عدد قليل من مرات القياس، غير أن هذا لا يقلل من منطق توزيع تباين الخطأ على جانبي المنحنى الاعتدالي.

الثانية: أن الدرجة الحقيقية للفرد وفقاً لهذا المنطق هي الأكثر استقراراً بينما

التغيير يحدث في الجزء الزائف من الدرجة الكلية، وهذا يعني بتعبير آخر نتيجتين محدودتين هما:

1- أننا لا نتوقع ارتباطاً بين الدرجة الحقيقية والدرجة الزائفة أو التباين الحقيقي وتباين الخطأ.

2- فأحدهما متغير والأخر ثابت.

3- أننا لا نتوقع أيضاً وجود ارتباط بين الدرجات الخاطئة في صورة ما من اختبار والدرجات الخاطئة في صورة متكافئة لنفس الاختبار، لأن الدرجات الزائفة عشوائية الاتجاه وغير منتظمة ولا يحكمها منطق القدرة الحقيقية التي يقيسها الاختبار.

وهذا المنطق يؤدي إلى صياغة عدد من المعادلات الجديدة التي تترتب على معادلة (1) التي مر ذكرها، والتي توضح هذه المعادلات طبيعة تباين الخطأ والتباين الحقيقي:

معادلة (2): متوسط تباين الخطأ = صفر

لأن تباين الخطأ الموجب والسالب يتوزع اعتدالياً.

معادلة (3): الارتباط بين التباين الحقيقي وتباين الخطأ = صفر

على افتراض استقرار الأداء الحقيقي وتذبذب تباين الخطأ إيجاباً وسلباً بين المرة والأخرى من مرات القياس.

معادلة (4): الارتباط بين الدرجات الخاطئة لصورتي نفس الاختبار = صفر

على افتراض عشوائية اتجاه تباين الخطأ في كل مرة من مرات الاختبار وبين جزئي الاختبار الواحد.. وهي النتيجة التي تعتمد عليها كل أساليب التصنيف في حساب الثبات.

لقد اعتمدنا في ترتيب المعادلات السابقة من خلال موقف افترضنا فيه أننا نقوم باختبار الشخص الواحد عشرات المرات اختباراً مستقلاً وبنفس المقياس، غير أن

هذا الإجراء غير ممكن عملياً لأسباب كثيرة، ولذلك نعود مرة أخرى من المعادلة (1) ومن المعادلات الأخرى المتسقة معها لنضع عدداً من الصيغ الجبرية التي تتعلق بالمكونات الصحيحة والخاطئة للمتوسط والتباين الخاص بالدرجة التي نحصل عليها من المقياس.

وبما أن مجموع متوسطات أي عدد من المكونات الموحدة تساوي متوسط هذه الجاميع فيمكننا أن نضع المعادلة الآتية:

معادلة (5) متوسط الدرجات الكلية = متوسط التباين الحقيقي + متوسط تباين الخطأ

ولما كان متوسط تباين الخطأ = صفراً (معادلة 2) ولذلك فإن متوسط الدرجات الكلية سيساوي متوسط التباين الحقيقي. ولما كان منطقياً أن متوسط مجموعة التباينات يساوي مجموع متوسطات هذه المجموعة من التباينات فنستطيع أن نضع معادلة للتباين كما وضعت للمتوسط وهي

معادلة (6): التباين الكلي = التباين الحقيقي + تباين الخطأ

وبهذه المعادلة نكون قد وصلنا إلى عدد من المفاهيم التالية، التي تمكننا من الانتقال إلى الوسائل الإجرائية لاستخلاص الثبات أو التباين الحقيقي للاختبار، والمفاهيم هي:

1- أن متوسط تباين الخطأ يساوي صفر.
2- لا يوجد ارتباط بين المقادير الخاطئة.

ويذلك يكون الثبات هو "النسبة من التباين الصحيح في الدرجات المستخلصة من الاختبار" ويمكن أن نعبر عن هذه النسبة بالمعادلة الآتية :

معادلة .. (7)

$$\text{ثبات الاختبار} = \frac{\text{التباين الحقيقي}}{\text{التباين الكلي}}$$

وعلينا الآن أن نتعرف على حجم التباين الحقيقي لكي نحسب ثبات الدرجة على الاختبار، وهذا ممكن إذ أتضح من المعادلة (6) أن هناك ارتباط بين التباين الكلي ومكونات الاختبار (التباين الحقيقي وتباين الخطأ)

وهذا الارتباط يتيح لنا الحصول على صيغة جديدة للمعادلة (6) بالمعادلة الآتية:

معادلة (8) .... التباين الحقيقي = التباين الكلي - تباين الخطأ

وبالتعويض عن هذه الفروق في المعادلة (8) مع المعادلة (7)

نحصل على المعادلة (9)

معادلة.. (9)

$$\text{معامل الثبات} = 1 - \frac{\text{تباين الخطأ}}{\text{التباين الكلي}}$$

ومن مميزات هذه المعادلة (9) هي أن البيانات التجريبية توفر لنا المعلومات اللازمة لحلها، حيث نستطيع أن نحصل على حجم التباين الكلي وتباين الخطأ وبالتالي نحسب معامل الثبات.

حساب التباين الحقيقي:

يمكن من خلال صيغة المعادلة رقم (7) أن نحصل عملياً على تقدير للتباين الحقيقي لدرجات الاختبار، فإذا كان التباين الكلي ومعامل الثبات معروفين فتحل المعادلة (7)

$$\text{معامل الثبات} = \frac{\text{التباين الحقيقي}}{\text{التباين الكلي}}$$

تصبح لدينا معادلة رقم (10)

معادلة رقم (10) التباين الحقيقي = معامل الثبات × التباين الكلي.

وبإيجاد الجذر التربيعي للتباينات في المعادلة السابقة يصبح لدينا تقدير للانحراف المعياري للدرجات الصحيحة.

$$\text{الانحراف المعياري الحقيقي} = \sqrt{\text{الانحراف المعياري الكلي} \times \text{معامل الثبات}}$$

أما مصادر الخطأ التي تؤدي إلى أخطاء القياس تتمثل في الآتي:

1- مستوى الفرد في سمة عامة معينة يؤثر على أدائه لعدد من المقاييس.
2- مدى فهم تعليمات المقياس وطريقة أدائه.
3- دافعية الفرد للإجابة عن المقياس.
4- التزييف والكذب والخداع من قبل الفرد.
5- الظروف البيئية مثل التهوية ودرجة الحرارة والضوضاء.
6- تعود الفرد على أداء نوع معين من الاختبارات والمقاييس.
7- الاضطرابات العقلية البسيطة مثل اضطراب الذاكرة والانتباه.
8- دور عامل الصدفة والتخمين في اختيار الإجابات الصحيحة.

وإذا ما استطعنا التعرف على بعض هذه المصادر فإنه يمكن علاجها وبالتالي يزيد ثبات المقياس، حيث نقلل من تأثيرها ليبقى التباين الموجود بين درجات الأفراد يعكس الفروق الفردية الحقيقية في الصفة أو السمة المراد قياسها لدى هؤلاء الأفراد.

والخلاصة من كل ما سبق يمكن إيجازها في أن المقياس يزيد ثباته كلما تحققت النقاط التالية:

1- الحصول على نفس النتائج تقريباً عند كل مرة يتم فيها إعادة التطبيق.
2- إن يكون التباين الحقيقي اكبر ما يمكن بالنسبة للتباين العام أو أن تباين الخطأ اقل ما يمكن.

3- إن يكون هناك علاقة بينية وارتباطية بين وحدات وأجزاء المقياس. لان وجود مثل هذه العلاقة يعطي دلالة على تناسق البناء الداخلي للمقياس وبالتالي فإن معامل الثبات يعتمد على هذه العلاقة البينية بين كل وحدة وأخرى وقيمة الارتباط بين كل وحدة وبين المقياس ككل. حيث إن تماسك المقياس وتناسق بنائه يدل على ثباته بل إنه يمكن حساب معامل الثبات من خلال هذه العلاقة بين وحدات المقياس.

العوامل المؤثرة في ثبات الاختبار:

هناك عوامل عديدة يمكن أن تؤثر على ثبات الاختبار. وبالتالي ينبغي من الباحث أن يتعرف عليها ويعمل على التحكم بها قدر الإمكان وتخفيض تدخلها في الثبات، ومن هذه العوامل:

أولا: طول الاختبار (عدد الفقرات أو الأسئلة):

يؤثر طول الاختبار أو عدد فقراته في درجة ثباته، إذ كلما تزايد عدد الفقرات كلما أرتفع ثبات الاختبار، وذلك لأن العدد الأكبر من الفقرات يؤدي إلى الحصول على عينة أكبر من السلوك، وبالتالي يكون من المتوقع أن يمثل بهذه العينة بشكل مستقر العدد الأكبر من مكونات السلوك أو السمة أو القدرة المقاسة والقابلة للظهور في مرتي التطبيق أو في نصفي أو جزئي الاختبار. في حين أن الفقرات القليلة في المقياس تقلل من احتمال استقرار مكونات السلوك أو الصفة المراد قياسها، وبالتالي انخفاض معامل الثبات. وكلما كان الاختبار طويلاً كلما زادت فرصة أن الخطأ يلغي بعضه بعضاً، وبالتالي زادت قيمة التباين مما يؤدي إلى زيادة معامل الثبات.

إن تباين الخطأ ثابت والتغير في التباين الكلي يعود إلى تباين الدرجة الحقيقية، وبالتالي عندما يزيد تباين الدرجة الحقيقية فإن التباين الكلي يزداد، حيث ان زيادة عدد فقرات الاختبار تعطي فرصة اكبر لرصد تباين الدرجة الحقيقية وبالتالي زيادة

التباين الكلي، وهو ما تؤكده الحقيقة التالية: (إذا زاد طول الاختبار (ن) مرة فإن التباين الحقيقي لدرجته يزيد (ن) 2 مرة ويزيد تباين الخطأ (ن) مرة، وبالتالي زيادة التباين الكلي يؤدي إلى معامل ثبات اكبر). وللوصول إلى معامل الثبات المرغوب للاختبار ينبغي أضافه عدد آخر من الفقرات إلي الاختبار، ويمكن استخدام معادلة (سبيرمان – بروان) في تحديد عدد الفقرات التي ينبغي أضافتها للوصول إلى معامل الثبات المرغوب. كما يمكن استخدام هذه المعادلة في تقليل عدد الفقرات بمعامل ثبات معين، حيث أن معادلة (سبيرمان – براون) تساعد على التنبؤ بالزيادة المتظرة في معامل الثبات إذا أضفنا إلى الاختبار عدداً من الفقرات المشابهة للفقرات الموجودة فيه، وبعبارة أخرى تمكن هذه المعادلة من تحديد العلاقة بين الزيادة في طول الاختبار والزيادة في معامل الاختبار مثلاً، إذا زاد عدد الفقرات من 25 إلى 75 فقرة، فعندها يصبح عدد فقرات الاختبار (3) مرات عددها قبل الزيادة، فتصبح ن = 3 أما إذا اختصرنا الأسئلة أو الفقرات من (100) فقرة إلى (50) فقرة، فتصبح ن = ½، ومن ثم نستخدم معادلة (سبيرمان – براون) التالية:

$$ر_ن = \frac{ن \times ر_س}{(ن - 1)\ ر_س + ن}$$

حيث أن:

$ر_ن$: معامل الثبات التقديري بعد زيادة فقرات الاختبار.

ن: عدد مرات إطالة الاختبار أو اختصاره

$ر_س$: معامل الثبات الذي حصلنا عليه قبل الزيادة أو النقصان.

فإذا كان معامل الثبات قبل الزيادة (0,50) فإن تطبيق المعادلة يكون بالشكل الأتي:

$$ر_ن = \frac{(0,50)\ (3)}{1 + (0,50)\ (1-3)}$$

وفي حالة نقصان الفقرات كما في المثال السابق وكان لدينا معامل الثبات (6,80) فإن تطبيق المعادلة يكون بالشكل الأتي:

$$ر_ن = \frac{(0,80)\ (0,50)}{1 + (0,80)\ (1 - 0,50)}$$

ويمكن اختصار معادلة (سبيرمان - براون) السابقة فيما إذا أردنا أن نتنبأ بمعامل ثبات اختبار يبلغ طوله ضعف طوله الأول كما هو الحال في حساب الثبات بالتجزئة النصفية ولذلك تصبح معادلة (سبيرمان - براون) كما يأتي:

$$ر_2 = \frac{2\ ر_2}{1 + ر_س}$$

حيث أن:

$ر_2$: ثبات الاختبار المصحح لنصفي الاختبار.

2: عدد مرات الإطالة للاختبار - كما في التجزئة النصفية إلى مرتين.

ر س: معامل الثبات قبل التصحيح بطريقة معامل الارتباط.

لكن ينبغي أن نشير هنا أن تطويل الاختبار ليس هو فقط الطريقة الوحيدة المفضلة لتحسين ثباته إذ أن المعرفة الصحيحة والتكنيك الدقيق للاختبار سيعطي معامل ثبات عالي.

ثانيا: مدى تقارب صعوبة الفقرات ودقتها:

أن وجود فقرات شديدة الصعوبة في الاختبار لا يستطيع جميع الأفراد أو أغلبهم الإجابة عنها لا تضيف ميزة للاختبار ولا يؤثر حذفها في الدرجة التي يحصل عليها أي فرد من أفراد العينة ما دامت لا تميز بين فرداً وأخر، وبالمثل فإن وجود فقرات شديدة السهولة التي يستطيع الإجابة عنها كل أفراد العينة فأنها لا تميز بين فرد" وأخر، لذلك ينبغي حذف أو تعديل الفقرات الشديدة الصعوبة أو السهولة إذا أردنا أن نرفع ثبات الاختبار لأنها لا تضيف شيئاً لقيمة درجة الثبات، لأن معامل

الثبات مبني في الأساس على التباين الحقيقي بين درجات أفراد العينة. وعليه فإن أفضل فقرات في الاختبار هي التي يكون احتمال الإجابة عليها 50 ٪ من الأفراد، كما أن الفقرات التي قد تحتمل تفسيرات عديدة ومتشعبة فأنها ستؤدي إلى انخفاض تباين الأسئلة وبالتالي انخفاض التباين الكلي للاختبار.

القيمة المقبولة لمعامل الثبات:

معامل الثبات عبارة عن نسبة تباين، ومعامل الثبات المقبولة تختلف من اختبار لآخر وذلك حسب الغرض من الاختبار ودقة القرار المترتب عليه وعلى ذلك:

1- الاختبارات المقننة تتطلب معاملات ثبات لا تقل عن (0,80) وذلك عندما تكون القرارات على مستوى الأفراد، أما على مستوى الجماعات فيقبل معامل ثبات (0,65).

2- معاملات ثبات الاختبارات التحصيلية المقننة يجب إلا تقل معاملات ثباتها عن (0,85) بينما يمكن أن تقل عن ذلك في اختبارات الشخصية.

3- الاختبارات التحصيلية والتي لا يراعى في تطبيقها الخطوات التي يتم مراعاتها في الاختبارات المقننة معظمها تتراوح معاملات ثباتها بين (0,20 – 0,40) ونادرا ما تصل إلى (0,60).

ثالثا: تعلق الإجابة على فقرة بالإجابة على فقرات أخرى:

عند استخدام اختبار يتضمن عدد من الفقرات ترتبط الإجابة عليها بعضها ببعضها الأخر يؤدي إلى انخفاض معامل ثبات الاختبار، لأن هذه الخاصية ستؤدي إلى خفض عدد الفقرات، مما ينبغي أن تكون كل فقرة مستقلة في نوعية الإجابة عليها حتى وأن كانت تقيس سمة واحدة أو قدرة واحدة.

رابعا: موضوعية التصحيح:

تعتبر موضوعية التصحيح من العوامل المؤثرة في ثبات الاختبار، ولاسيما في

الاختبارات التي تعتمد على تقدير المصحح كاختبارات المقال التحصيلية واختبارات الإبداع أو الاختبارات الإسقاطية، حيث أن تباين التصحيح يؤدي إلى زيادة تباين الخطأ وبالتالي إلى نقصان معامل الثبات. مما ينبغي اللجوء إلى تقدير معامل ثبات المصححين، كما ينبغي أن يتضمن الاختبار وصفاً دقيقاً لمحكات الإجابة والتصحيح وطريقة وضع الدرجة عليها.

خامسا: أثر تخمين المجيب:

في الاختبارات التي تعتمد على تقدير الفقرة أو السؤال بعدد من البدائل، يلجأ بعض المفحوصين عادة في حالة عدم تأكدهم من الإجابة الصحيحة إلى التخمين واختيار أي أجابه منها، وعادة ما يكون التخمين في مثل ذلك عند أعاده الاختبار أو في جزء آخر من الاختبار، مما يؤدي إلى خفض ثبات الاختبار ويمكن معالجة أثر التخمين أو الكذب لاسيما في اختبارات الشخصية في إحدى الطرق التالية:

1- من الممكن أضافه بعض الفقرات الكاشفة للكذب بين فقرات الاختبار أو المقياس والتي من خلالها نتمكن من معرفة دقة الإجابات وعدم الكذب في الإجابة وبالتالي يتمكن الباحث من إهمال اختبار الفرد الذي تكون درجة الكذب لديه عالية في المقياس أو الاختبار. ويستخدم هذا الأسلوب عادة في اختبارات أو مقاييس الشخصية واختبارات القدرات المميزة.

2- وضع بعض الضوابط التي تجبر المفحوص على التفكير على الإجابة قبل أن يؤشر عليها كأن تكن الفقرة مكونة من عبارتين لهما نفس المقبولية الاجتماعية لكنهما يختلفان في قياس الظاهرة، ويحدث هذا في اختبارات القدرات المميزة أيضاً.

3- أما في الاختبارات التحصيلية فيمكن استخدام المعادلة الآتية بمعالجة أثر التخمين:

$$\text{د ح} = \frac{\text{د ص} - \text{د خ}}{\text{ن} - 1}$$

حيث أن:

د ح = الدرجة الحقيقة

د ص = عدد الفقرات الصحيحة

د خ = عدد الفقرات الخاطئة

ن = عدد بدائل الإجابة على الفقرة

سادسا: زمن الاختبار:

يؤثر الزمن المحدد للإجابة على الاختبار على ثباته، فيزداد الثبات تبعاً لزيادة الزمن حتى يصل إلى الحد المناسب للاختبار - أي الزمن الذي تتطلبه الإجابة على الاختبار - فيصل الثبات إلى أعلى درجة ممكنة من جراء تأثير هذا العامل. أما إذا كان الاختبار يتطلب وقتاً كبيراً لإنجازه، فإن ذلك سيؤدي إلى خفض معامل الثبات، لأن ذلك قد يؤدي إلى الإجهاد والتعب وضعف الدافعية على الإجابة وبالتالي قد تتعرض الفقرات الأخيرة من الاختبار إلى أخطاء في الإجابة ومن ثم تؤدي إلى زيادة تباين الخطأ.

وعليه ينبغي تحديد الزمن المطلوب للاختبار ووفقاً للجهد المطلوب للإجابة عليه ومدى قدرة المفحوص على الاستمرار في الإجابة دون تعب، وتحديد مدة الاختبار المطلوبة يمكن الوصول إليها عن طريق تجريبه على عينات استطلاعية

سابعا: تجانس العينة:

يؤدي التجانس الشديد في عينة الثبات من حيث الظاهرة المدروسة إلى انخفاض واضح في معامل الثبات، لأن التباين داخل هذه العينة المتجانسة يكون منخفضاً بقدر لا يسمح بتقدير التباين الحقيقي - أي ثباته - في حين كلما كبر حجم العينة وتزايد عدم تجانسها كلما كانت أقرب إلى التوزيع الاعتدالي وبالتالي سيصبح تباين الخطأ صغيراً مما تؤدي هذه النتيجة إلى ارتفاع ثبات الاختبار.

ويمكن تفسير ذلك أيضاً إلى أن تجانس أفراد العينة سوف يؤدي إلى انخفاض تباين كل سؤال، حيث أن تباين السؤال يعتمد على نسبة الأفراد الذين أجابوا إجابة صحيحة ونسبة الأفراد الذين أجابوا إجابة خاطئة، وبالتالي انخفاض تباين الأسئلة يؤدي إلى انخفاض التباين الكلي للاختبار، ومن المعلوم أن تباين الخطأ ثابت وبالتالي فإن قسمة تباين الخطأ على التباين الكلي سوف يكون كبيرا وعند طرحه من الواحد الصحيح للحصول على معامل الثبات فإن معامل الثبات يكون منخفضاً.

ثامنا: بعض العوامل الأخرى التي تؤثر بمقادير مختلفة في ثبات الاختبار:

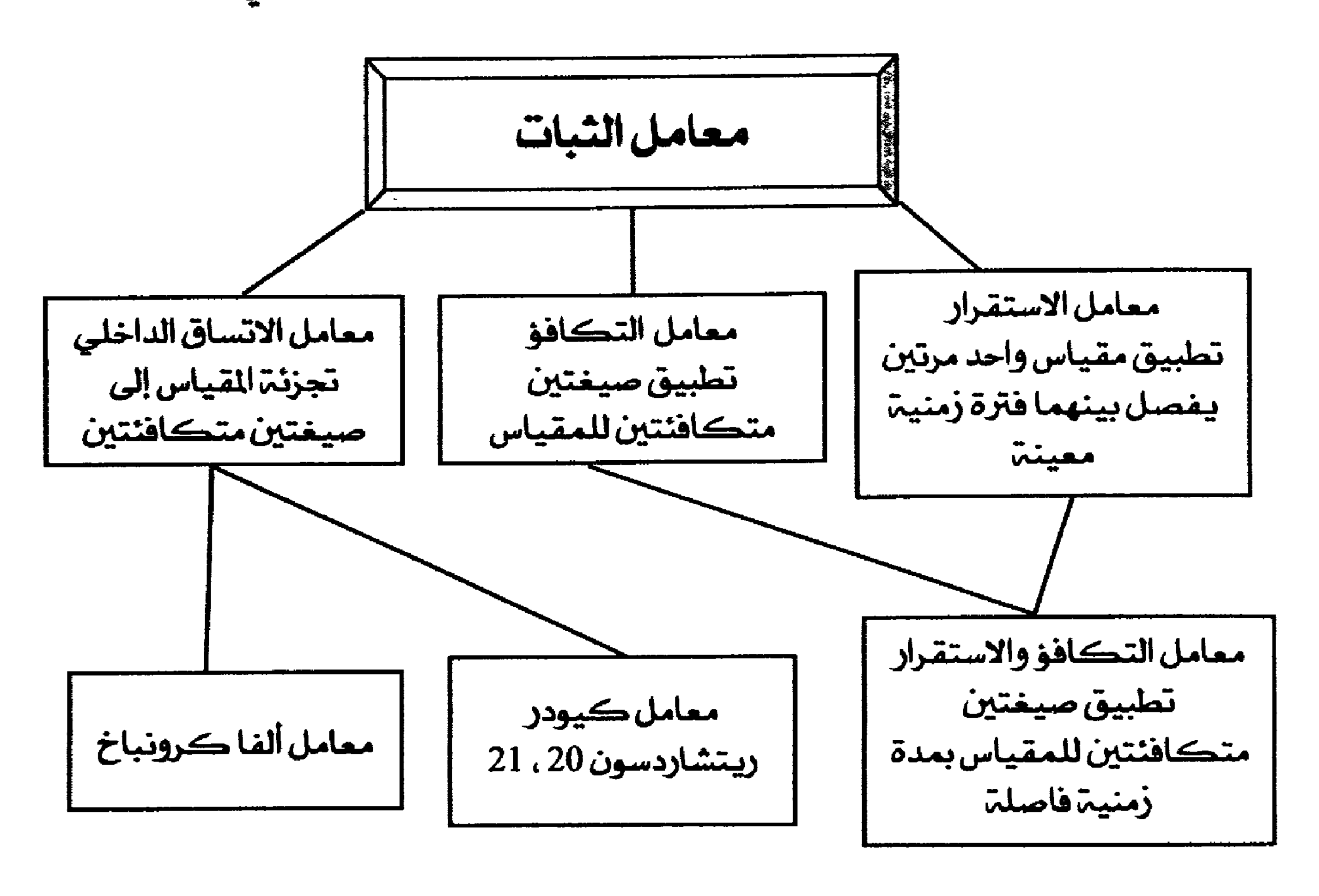
كضبط موقف التطبيق ودافعيه المفحوص والمؤثرات الفيزيقية والمشتتات المتعددة في موقف الاختبار والحالة الصحية والانفعالية للمفحوص مما ينبغي أن يقوم الباحث بضبط دقيق لكل هذه العوامل حتى يتمكن من الوصول إلى معاملات ثبات مرتفعة.

أساليب حساب معامل الثبات:

تتعدد وتتنوع أساليب حساب الثبات في الاختبارات والمقاييس، ويختص كل أسلوب منها بتقدير نوعية محددة من (تباين الخطأ) وهو التباين الذي يؤثر في ثبات القياس الذي نحصل عليه كلما استخدمنا مقاييس مختلفة.

وتتميز بعض الاختبارات والمقاييس بإمكان حساب ثباتها بأسلوب دون آخر وعدم صلاحية بعض الأساليب لها، إلا أن هذا لا يعني أن المقياس أو الاختبار الواحد لا يصلح له إلا أسلوب واحد من أساليب حساب الثبات حيث أن في بعض الحالات يمكن أن تستخدم أكثر من أسلوب واحد للاختبار الواحد بهدف التعرف على مصادر تباين الخطأ الذي يؤثر في استقرار أو أتساق الدرجة التي نحصل عليها من الاختبار.

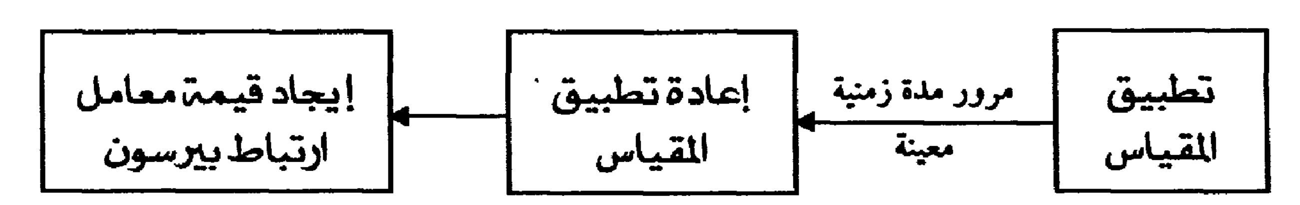
وفيما يلي استعراض موجز لأساليب حساب الثبات ونوعية الثبات التي يقيسها كل أسلوب وطريقة حسابه ويمكن تلخيص هذه الطرق بالمخطط الاتي:



أولا: إعادة الاختبار: Test – retest

يسمى معامل الثبات الناتج من هذه الطريقة بمعامل الاستقرار coefficient of stability، أي استقرار نتائج الاختبار خلال الفترة بين التطبيق الأول والثاني للاختبار. حيث أن هذا النوع في حقيقته هو معامل الارتباط بين مجموعتي درجات الاختبار في التطبيقين، وذلك بإعطاء نفس مجموعة الأفراد الاختبار نفسه مرتين، ومصدر الخطأ في الدرجة التي يحصل عليها الفرد في ضوء معامل الثبات هو خطأ التغيرات عبر الفقرة الزمنية. ولا يستخدم هذا الأسلوب في استخراج الاتساق الداخلي، إذ قد يكون معامل الارتباط عالياً بينما يكون معامل الثبات الداخلي منخفضاً. ويفضل عادة عند حساب الثبات بهذا الأسلوب الا يكتفي الباحث

بحسابه على مدى فترة زمنية واحدة بل على أكثر من فترة، إذ من الممكن أن يعاد التطبيق بعد أسبوعين ومن ثم أعادته بعد فترة شهر ثم أعادته مرة ثالثة بعد فترة أخرى. ويستفاد من هذه النتائج في إمكانية الباحث بالتنبؤ بنتائج اختباره عبر الزمن، كما يمكن للباحث أن يستخدم أكثر من عينة متجانسة تماماً ويعيد عليها الاختبار بعد فترة مختلفة كي يتخلص تقريباً من احتمال تأثر أداء العينة بإعادة الاختبار سواء كان هذا التأثر إيجابياً أو سلبياً نتيجة للمران أو التدريب والنمو أو نتيجة للملل أو عدم الاهتمام.



ويواجه هذا الأسلوب (أعاده الاختبار) في إيجاد معامل الثبات بعض الانتقادات والعيوب، لعل من أهمهما ما يأتي:

1- تمثل الفقرات التي استخدمت في الاختبارين عينة واحدة فقط. والتي تكون عادة من مجتمع واحد من الفقرات المتيسرة للاختبار، مما لا تكون درجات أعاده الاختبار دليلاً واضحاً عن كمية تغير مقدار الدرجات فيما إذا استعملنا عينة مختلفة من الفقرات أو الأسئلة.

2- إن إجابات المفحوص في الاختبار الثاني لا تكون مستقلة عن إجاباته على الاختبار الأول، إذ قد تتأثر استجاباتهم في المرة الثانية في تذكرهم لاستجاباتهم في التطبيق الأول، أو من خلال المناقشات التي تجرى بين المفحوصين بعد الانتهاء من التطبيق الأول، أو قد تتأثر بواسطة أنواع أخرى من الجهود التي يبذلها بعض المفحوصين عادة لمعرفة الإجابة الصحيحة بعد الانتهاء من التطبيق الأول مما تساعدهم هذه الجهود في معالجة الصعوبات التي واجهوها وبالتالي ستتحسن إجاباتهم في التطبيق الثاني.

3- قد تتداخل أخطاء القياس مع التغيرات الحقيقية في قابليات وقدرات المفحوصين عند أعاده الاختبار نتيجة للتعلم أو النمو.

4- في أحيان كثيرة بعض المفحوصين لاسيما من الطلبة لا يروق لهم استخدام هذا الأسلوب باعتباره مضيعة لوقت التعلم وإرباك للنظام المدرسي لاسيما في التطبيق الثاني، كما أن المعلمين لا يحبذون ضياع وقت الطلبة في مثل هذه الأمور، مما يؤدي ذلك إلى ضعف رغبة قسم من الطلبة في الإجابة على الاختبار الثاني أكثر من الأول وبذلك يقل ثبات الاختبار.

5- التذبذب العشوائي في أداء الفرد الواحد في الموقعين، لأن أداء الفرد لا يتم من خلال قوالب صارمة في كل الأوقات، وبالإضافة إلى التغيرات التي قد تحدث في حالته الصحية أو النفسية أو تغيرات الجو وما إلى ذلك من مؤثرات على الفرد أثناء الاختبار مما تؤثر بشكل ما على معامل الثبات.

ويمكن للباحث أن يقلل من مساوئ أعاده الاختبار عن طريق أعداد صورتين متكافئتين من اختبار معين، تعطى الصورة الأولى في وقت معين وتعطي الثانية بعد ذلك فترة، ثم يحسب معامل الارتباط بين درجات الصورتين لإيجاد معامل الثبات وبذلك يتأكد الباحث بهذه الطريقة من استقرار الاختبار وتكافئه.

شروط استخدام أسلوب (أعاده الاختبار) في إيجاد الثبات:

إن استخدام هذا الأسلوب في إيجاد ثبات الاختبار أو المقياس محكوم ببعض الشروط والقواعد التي ينبغي على الباحث مراعاتها والتأكد منها قبل استخدام أسلوب (أعاده الاختبار) في إيجاد معامل الثبات، لعل من أهمها ما يأتي:

أولاً: أن تتسم الظاهرة المدروسة بنوع من الاستقرار، وليست متذبذبة.

ثانياً: أن يبتعد الباحث عن استخدام هذا الأسلوب في الاختبارات التي قد يؤدي التطبيق الأول لها إلى نوع من التعلم أو التدريب على الاستجابات الصحيحة، أو التي تخضع إلى التذكر، لذلك فإن هذا الأسلوب لا يصلح للاختبارات والمقاييس

التي تهدف إلى قياس التذكر. أو التي تؤدي الإجابة عليها إلى نوع من التدريب والتعلم، كاختبارات التحصيل وبعض الاختبارات التي ترتبط بالعمليات العقلية، في حين يمكن استخدامه في اختبارات الشخصية بعد مراعاة الفاصل الزمني المفضل بين التطبيقين، وتصلح للاختبارات ذات الزمن المحدد إلى حد كبير، والاختبارات ذات الزمن غير المحدد لكونه يقيس قوة الاستجابة أكثر من سرعتها.

ثالثاً: أن يعمل الباحث على توفير نفس الظروف للمفحوصين في الموقف الاختباري، الأول والثاني، لأن الاستجابات قد تتأثر بالظروف المحيطة بالمفحوصين أكثر مما تتعلق بالمقياس أو الاختبار.

رابعاً: على الباحث مراعاة الفاصل الزمني بين التطبيق الأول والثاني للاختبار، والتي لا يمكن تحديد فترة ملائمة لجميع المقاييس والاختبارات، إذ أن طول هذه الفترة يعتمد عموماً على طبيعة الظاهرة المدروسة وطبيعة عينة البحث حيث أن درجة الثبات أو معامل الاستقرار يتأثر بطول الفترة الزمنية بين التطبيقين للاختبار، فإذا كانت قصيرة جداً فهناك احتمال كبير أن يتذكر المجيب إجاباته في التطبيق الأول، في حين إذا كانت الفترة طويلة جداً فقد يحصل تغير حقيقي في الظاهرة أو الصفة المدروسة مما يرتفع الثبات في الحالة الأولى ويقل في الحالة الثانية.

لذلك يفضل أن تكون الفترة الزمنية قصيرة في الحالات التالية:

أ- إذا كانت الظاهرة تتأثر بالنمو الجسمي والعقلي لأفراد العينة مثلاً وفي الممارسة العملية يتطلب أعاده الاختبار في إيجاد الثبات لها فترة قصيرة بعد التطبيق الأول للاختبار.

ب- عند اختبار الأطفال الصغار وبسبب سرعة معدل النمو لديهم، لأن التغيرات التي قد تحدث في الأداء النسبي بعد فترة زمنية طويلة يمكن اعتبارها ضرباً من النمو والتقدم وليس مجرد تغيير عشوائي ناتج عن الظروف التي قد تطرأ على الموقف الاختياري التي لا يمكن التحكم فيها.

ج- في الاختبارات التي لا تتأثر الاستجابات في التطبيق الثاني بتذكر استجابات التطبيق الأول لها.

في حين يفضل أن تكون الفترة الزمنية طويلة في الحالات التالية:

أ - إذا كانت الظاهرة المدروسة لا تتأثر بالنمو والتغيير السريع.

ب- في حالة كون أفراد العينة كبار السن ولديهم نوعاً من استقرار النمو.

ج- إذا كانت الاستجابات تتأثر بالتذكر من التطبيق الأول للاختبار.

وعموماً فإن تحديد الفترة الفاصلة بين التطبيقين للاختبار يتوقف على دراسة طبيعة الظاهرة وعينة التطبيق، وقد تمتد هذه الفترة ما بين نصف ساعة إلى ستة أشهر وأحياناً أكثر، في حين يرى البعض أن لا تقل هذه الفترة عن أسبوع ولا تزيد عن ستة أشهر، وكما سبق أن بينا أن تحديد الفترة من قبل الباحث لا يتم اعتباطاً بل حسب طبيعة البحث ونوعية العينة وخصائص الظاهرة المدروسة.

خامساً: أما كيفية حساب معامل الثبات بهذا الأسلوب فيتم بالخطوات الآتية:

أ - يقدم الاختبار أو المقياس إلى أفراد العينة ثم يصحح وتدون نتيجة أو درجة كل فرد.

ب- يعاد تطبيق الاختبار على أفراد العينة بعد الفاصل الزمني الملائم وضمن ظروف مشابهة في التطبيق الأول، ثم يصحح حسب قواعد تصحيح التطبيق الأول له ثم تدون النتائج أو الدرجات لكل فرد أمام درجته في التطبيق الأول.

ج- يتم حساب معامل الارتباط بين مجموعتي الدرجات - درجات التطبيق الأول ودرجات التطبيق الثاني - ومعامل الارتباط هذا يمثل معامل الثبات، وعلى الباحث أن يراعي طبيعة البيانات عند استخدام معامل الارتباط، فإذا كانت متصلة أي غير منفصلة فيمكن استخدام (معامل ارتباط - بيرسون) أما إذا كانت البيانات من النوع المنفصل فيفضل استخدام (معامل ارتباط - سبيرمان للرتب) إلا إنه في حالة وجود درجات أو قيم مماثلة كثيرة، أي أن نفس الدرجة

تتكرر لدى عدد من الأفراد فإنه لا يفضل استخدام (معامل ارتباط سبيرمان) بل يفضل استخدام (معامل ارتباط كاندل).

وتعد هذه الطريقة مناسبة لاختبارات الاستعداد وذلك لان نتائج اختبار الاستعداد تستخدم في التنبؤ بعيد المدى عن مستقبل الطالب التربوي أو المهني، وبالتالي فإنه من الضروري أن نتعرف على مدى استقرار درجات الاستعداد، كما تعتبر مناسبة لحساب معامل ثبات الاختبارات الموقوتة ذات الزمن المحدد والتي تعتمد إلى حد كبير على السرعة، كما أنها تلائم مقاييس الاتجاهات والميول اعتمادا على أن الاتجاهات والميول بطيئة التغيير، ونجد أن هذه الطريقة لا تصلح للاختبارات التي تهدف إلى قياس التذكر أو ترتبط به ارتباطاً مباشراً، كما أنها لا تناسب اختبارات التحصيل لأن التحصيل سريع التغيير.

ثانيا: أسلوب الصور المتكافئة Equivalent forms

إذا أعددت صورتان متماثلتان أو أكثر لاختبار واحد بنفس الطريقة فمن المحتمل أن درجات هذه الصور البديلة سوف تظهر متكافئة، وحينها يتم إعطاء كل فرد من أفراد العينة صورتي الاختبار، فإن الارتباط بين الصورتين سيكون معاملا للثبات. ويسمى معامل الثبات بهذا الأسلوب بمعامل التكافؤ coefficient of equivalence حيث يدل على مدى الارتباط بين درجات اختبار واختبار آخر متكافئ معه يعطى في الوقت نفسه أو من وقت متقارب جداً وتباين الخطأ في هذا الأسلوب يدل على أخطاء عينة المحتوى – الأسئلة أو الفقرات – ويعد هذا الأسلوب من أفضل الأساليب في حساب الثبات في الاختبارات التحصيلية في حين لا يصلح للقياس في الاختبارات أو المقاييس الموقوتة.

ويمتاز أسلوب الصور المتكافئة ببعض المميزات التي تجعله أحياناً مفضلاً في حساب معامل الثبات ومن هذه المميزات ما يأتي:

1- يساعد هذا الأسلوب في خفض احتمال تأثير التدريب أو التعلم والتذكر على النتائج كما في أسلوب (إعادة الاختبار).

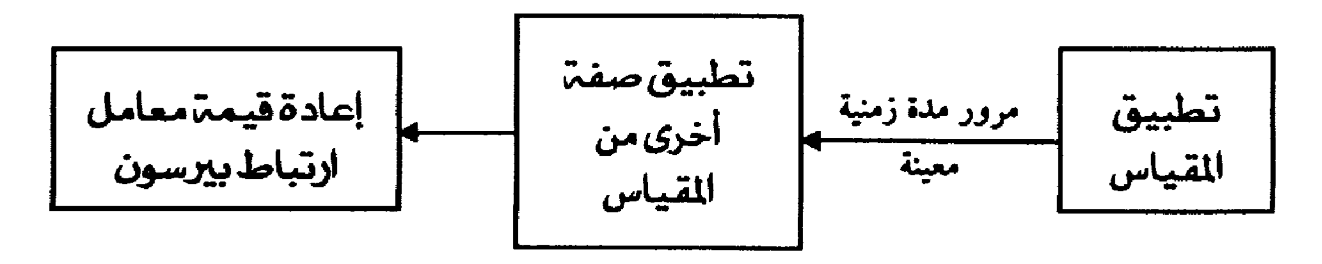
2- إذا تم استخدام الصورتين المتكافئتين في وقتين مختلفين، فأننا نضمن تلافي بعض العيوب (إعادة الاختبار) كما يمكن أن تقيس به معامل الاستقرار والتكافؤ معاً.

3- معامل الثبات باستخدام الصور المتكافئة فيه تشابه كبير مع معامل الاتساق الداخلي الذي يتم حسابه بأسلوب التجزئة النصفية، لاسيما في تطبيق الصورتين في الوقت نفسه أو متقاربة جداً.

4- تعد الصور المتكافئة من الأساليب المناسبة لتقدير مدى تمثيل عينة الفقرات للمجال الذي يقيسه الاختبار، فكلما كان هذا التمثيل جيداً، فأننا نتوقع أن ينعكس على الصورة المكافئة، وهننا يظهر أيضاً التجانس في أداء الفرد على الصورتين في شكل معامل الثبات.

إلا أن هذه الميزات لا تمنع من وجود بعض الصعوبات في هذا الأسلوب منها، صعوبة أعداد صورتين متكافئتين تماماً لها نفس المحتوى ونفس الهدف فيما تقيسه من الظاهرة المدروسة وما تحتاج من جهد ووقت وتكلفة عند إعدادهما.

ويمكن تلخيص طريقة حساب معامل التكافؤ بالمخطط الآتي:



قواعد وشروط أعداد الصور المتكافئة:

1- أن يبذل الباحث جهداً واضحاً للوصول إلى فقرات متكافئة في صور الاختبار سواء أكان ذلك في صياغتها أو مستوى صعوبتها أو تمثيلها للسمة المقاسة مع

استخدام نفس التعليمات، لكنها ذات استقلالية أي أنها غير متداخلة أو متشابهة في بعض الفقرات.

2- ويشير (ثورندايك) إلى أن الصور المتكافئة ينبغي أن يكون لها نفس التباين الحقيقي، كما ينبغي أن لا يحدث تداخلاً في تباين الخطأ فيها، إذ كلما كان الفرق بين التباين الحقيقي كبيراً كلما صغر معامل الثبات، وكلما تداخل تباين الخطأ كلما زاد معامل الثبات، لأن معامل الارتباط الذي يحسب بينهما يعتمد على تشابه التباين الحقيقي وتباين الخطأ التجريبي ويرى (جلكسون) أن الصور المتكافئة لها نفس المتوسط ونفس التباين، كما يفضل استخدام ثلاث صور للاختبار ومن ثم حساب معاملات الارتباط بينها.

3- عدد استخدام معامل الارتباط بين درجات العينة لكلا الصورتين، على الباحث أن يراعي طبيعة الدرجات للاستخدام نوع معامل الارتباط المناسب - كما سبق توضيحه في أسلوب أعاده الاختبار.

ثالثا: أسلوب التجزئة النصفية Spilt – halves:

يعتمد هذا الأسلوب أساساً على تقسيم فقرات المقياس أو الاختبار إلى قسمين متكافئين، ثم حساب معامل الارتباط بين درجات هذين القسمين ولذلك فإن هذا الأسلوب يصلح في الاختبارات التي تكون فيها فقرات المقياس أو الاختبار متجانسة أي أنها تقيس خاصية أو سمة واحدة وعليه فإنه يقيس التجانس بين الفقرات، أي الاتساق الداخلي، مما يسمى معامل الثبات بهذا الأسلوب (معامل الاتساق الداخلي) coefficient of consistency وأن تباين الخطأ الذي يقيسه أسلوب التجزئة النصفية هو عدم تجانس نصفي الاختبار.

ومن الواضح أن حساب الثبات بالتجزئة النصفية لا يقيس التجانس الكلي للمقياس لأنه يقسم الفقرات إلى قسمين، لذلك فهو معامل ثبات لنصف الاختبار أو المقياس، ولتلافي ذلك ينبغي تصحيح معامل الثبات المستخرج بطريقة حساب

الارتباط بين درجاتها باستخدام بعض الطرق الإحصائية لإيجاد معامل الثبات لكل الاختبار، ومنها معادلة (سبيرمان - براون) Spearmen – Brown formula كما هناك بعض الطرق التي يحسب بها كل معامل الثبات والتي سيمر ذكرها فيما بعد.

كما أن هذا الأسلوب يجنب الباحث من بعض الصعوبات التي تواجهه في إيجاد معامل الثبات بأسلوب (أعاده الاختبار) وبأسلوب (الصور المتكافئة) ويمكن اعتباره من أكثر الأساليب عملية وسهولة. كما ينبغي أن نؤكد أن تقسيم الاختبار إلى جزأين لا يتم قبل تطبيقه ولا يؤثر عليه، بل يقدم على شكل متكامل، ومن ثم يجري تقسيمه بعد ذلك، وهناك عدة طرق لتقسيم الاختبار أو المقياس إلى قسمين، من أهمها الآتي:

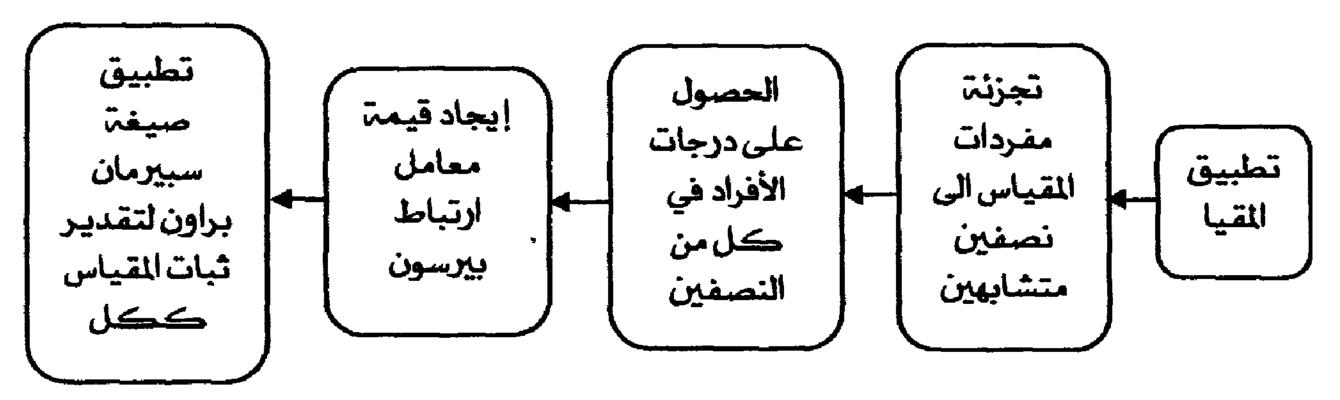
1- القسمة النصفية، وذلك بتقسيم فقرات الاختبار إلى نصفين متساويين يمثل النصف الأول المجموعة الأولى من الفقرات ويمثل النصف الثاني المجموعة الثانية من الفقرات، بيد إنه لا يفضل استخدام هذه الطريقة لاعتبارات عديدة منها أن المجيب قد لا يكون بنفس الدافعية أو القدرة عند الإجابة على النصف الثاني من الاختبار.

2- الفقرات الفردية والفقرات الزوجية، وهذه من الطرق الشائعة في التجزئة النصفية.

3- جزءا الاختبار: رغم أن طريقة التصنيف إلى فردي وزوجي من أفضل الطرق التي قد تؤدي إلى نصفين متعادلين في خصائصها ومتساويين في تعرضها لظروف الأداء، إلا أن بعض المقاييس والاختبارات لا تصلح هذه الطريقة في تقدير ثباتها لاسيما إذا كانت تقدم كل مجموعة من الفقرات في وحدة زمنية معينة أو يكون المطلوب من المبحوثين تقديم عدد كبير من الاستجابات لفترة واحدة في فترة زمنية معينة ثم تقدم الفقرة الأخرى وهكذا، ولذلك يفضل تقسيم الفقرات إلى (4) أجزاء مستقلة كل جزأين في قسم واحد، وقد يتكون القسم الأول من الجزأين (1، 3) والقسم الثاني من الجزأين (2، 4).

4- وهناك من يقسم الفقرات إلى قسمين بعد تحديد مستوى صعوبة كل فقرة ومن ثم توزيعها بالتساوي على قسمي الاختبار، كما يمكن توزيعها بناء أعلى مدى صلاحيتها للتمييز بين الأفراد، وهذا الإجراء يتطلب عادة تطبيق الاختبار على عينة استطلاعية بهدف تحديد مستوى صعوبة وتمييز كل فقرة.

ويمكن توضيح ذلك بالمخطط الآتي:



شروط استخدام أسلوب التجزئة النصفية:

1- كما يشير (جلكسن) إنه من الخطأ اعتبار التجزئة النصفية على أنها صورتان متكافئتان، ولكي تكون ذلك ينبغي أن تكون متعادلتين من حيث المتوسط ودرجة الصعوبة والتشتت ومعاملات الارتباط بين الوحدات.

2- لا تصلح إلا في اختبارات القوة لا السرعة، كي تتساوى الفقرات المستخدمة في حساب معامل الارتباط، ويكون هناك قدر كاف من الفقرات لحساب معامل الارتباط.

3- تستخدم في المقاييس والاختبارات التي تكون فيها الفقرات متجانسة أي أنها تقيس خاصية أو سمة واحدة.

طرق حساب الثبات لأسلوب التجزئة النصفية:

1- معادلة (سبيرمان - براون) (Spearman - Brown)

بما أن حساب الثبات بالتجزئة النصفية، هو عبارة عن ثبات نصف الاختبار

وليس كله، لذلك ينبغي تصحيح معامل الارتباط الذي يمثل معامل الثبات في التجزئة النصفية بمعادلة (سبيرمان - براون) حيث تساعد هذه المعادلة على حساب معامل الثبات إذا أردنا أن نضيف فقرات إلى الاختبار بعد حسابه بطريقة معامل الارتباط، أي أنها تساعد في تحديد العلاقة بين الزيادة في معامل الثبات وعدد الفقرات التي ينبغي أن تضاف إلى الاختبار.

بيد أن هناك بعض الانتقادات الموجهة لهذه المعادلة في تصحيح معامل الثبات الذي تم إيجاده بحساب معامل الارتباط، منها:

1- افتراضها تشتت واحد لنصف الاختبار، ونادراً ما يكون الانحراف المعياري متساوياً في النصفين.
2- تحتاج إلى عمليات حسابية كثيرة، لأنها تحتاج إلى حساب معامل الارتباط بين جزئي الاختبار.
3- مبالغتها في رفع قيمة معامل الثبات النصفي.
4- لا تصلح مع الاختبارات الموقوتة.

2. معادلة (رولون) Rulon:

تعتبر معادلة (رولون) طريقة مختصرة لحساب معامل الثبات الكلي للاختبار ولا تحتاج إلى إيجاد معامل الارتباط بين جزئي الاختبار كما في معادلة (سبيرمان - براون) وهي:

$$ر = 1 - \frac{ع_{د}^{2}}{ع_{س}^{2}}$$

حيث أن:

ر = معامل الثبات

$ع_{د}^{2}$ = تباين الفرق بين درجات نصفي الاختبار

$ع_{س}^{2}$ = التباين الكلي لمجموع درجات نصفي الاختبار

وتتجه معادلة (رولون) إلى تقدير الثبات بالاعتماد على نسبة التباين الحقيقية للأداء فقط مؤكدة أن الفرق بين التباين المنطقي الذي يفترض أن يعبر عنه الاختبار هو (1) وبين تباين الخطأ هو معامل الثبات الفعلي للاختبار. وعلى هذا فالمتوقع في حالة الثبات المطلق للاختبار إلا يوجد فرق بين أداء الفرد على نصفي الاختبار ولهذا تستخدم معادلة (رولون) هذا الفرق بين الدرجات على نصفي الاختبار لتقدير ثباته.

كما أن هذه المعادلة تحسب ثبات الاختبار كله دون حاجة لإعادة تصحيح الطول، ويرجع ذلك إلى حقيقة أن الفرق بين درجتي الفرد على نصفي الاختبار عبارة عن مجموع تباين الخطأ في النصفين معاً وليس نصفا واحداً فقط ولهذا فالفروق بين النصفين خاصة بالاختبار كله، أي أن التباين الذي تعبر عنه هذه الفروق هو الجزء من التباين في الدرجة الكلية الراجع للخطأ في كل الاختبار وليس تباين خطأ نصف الاختبار.

3- معادلة موزير Mosier:

تتميز معادلة موزير بأنها مختصرة، إذ أن توفر القدر الأكبر من الجهد المبذول في العمليات الحسابية، لكنها تتطلب التصحيح بمعادلة (سبيرمان - براون) ونص المعادلة كالأتي:

$$\text{ر ف ز} = \frac{(\text{ر ف ك})\,(\text{ع ك}) - \text{ع ف}}{\text{ع}^2_{\text{ك}} + \text{ع}^2_{\text{ف}} - (2)\,(\text{ر ف ك})\,(\text{ع ف})\,(\text{ع ك})}$$

حيث أن:

رف ز: الارتباط بين النصف الفردي والزوجي.

رف ك: الارتباط بين النصف الفردي مثلاً والاختبار كله.

ع ف: الانحراف المعياري لدرجات النصف الفردي.

ع ك: الانحراف المعياري لدرجات الاختبار كاملاً.

4- معادلة جتمان:

معادلة (جتمان) تعتمد على نفس المنطق الذي اعتمدت عليه معادلة (رولون) والفارق بينهما لا يخرج عن كونه فارقاً في أسلوب الحساب الذي يجعل من معادلة (جتمان) أكثر سهولة، إذ يستخدم فيها التباين الخاص بكلا النصفين حيث تقوم بجمع هذا التباين وقسمته على التباين الكلي للاختبار، كما لا تحتاج معادلة (جتمان) لاستخدام معادلة أخرى لتصحيح الطول، ونص المعادلة كالآتي:

$$ر_{ك} = 2\left(1 - \frac{ع_{1}^{2} + ع_{2}^{2}}{ع_{ك}^{2}}\right)$$

حيث أن:

$ر_{ك}$: ثبات الاختبار.

$ع_{1}^{2}$: تباين الجزء الأول من الاختبار.

$ع_{2}^{2}$: تباين الجزء الثاني من الاختبار.

$ع_{2}$ ك: تباين الاختبار كله.

5- معادلة هورست Horst:

عندما تمنع بعض الاعتبارات العملية أحياناً التصنيف المتوازن للاختبار كأن تقوم بقسمة الاختبار إلى جزئين غير متعادلين، مما لا يصح هنا استخدام أي من المعادلات السابقة ويمكن فقط استخدام معادلة (هورست)، ونص المعادلة الآتية:

$$ر_{ك} = \frac{\sqrt{[\,ر^{2} + 4\,(ك)\,(ص)\,(1 - ر^{2}) - ر\,]}}{2\,(ك)\,(ص)\,(1 - ر^{2})}$$

حيث أن:

$ر_{ك}$ = ثبات الاختبار كاملاً

ر = الارتباط بين جزئي الاختبار.

ص = النسبة الصغرى من الاختبار الممثلة للجزء الأصغر.

ك = النسبة الكبرى من الاختبار الممثلة للجزء الأكبر.

وتعد معادلة (هورست) معادلة تصحيح طول مثلها في ذلك مثل معادلة سبيرمان - براون، ما دامت تبدأ من حساب الارتباط بين جزأي الاختبار غير المتساويين وتنتهي إلى أعادة تقدير هذا الارتباط مدخلة في اعتبارها الطول المختلف لكل جزء.

6- معادلة جلكسون Guliksen:

معادلة التنبؤ (لسبيرمان - براون) تتأثر بالزمن المحدد للاختبار وبذلك لا تصلح مع حساب ثبات الاختبارات الموقوتة التي قد تحول دون إكمال الاختبار في الوقت المحدد للإجابة، لذلك أقترح (جلكسون) المعادلة الآتية لحساب ثبات الاختبارات الموقوتة:

$$ر''_{11} = ر_{11} - \frac{م\ ت}{ع^2\ خ}$$

حيث أن:

ر" 11 = معامل ثبات الاختبار الموقوت أو معامل الثبات بعد تصحيح أثر السرعة.

ر 11 = معامل الثبات المحسوب بطريقة (سبيرمان - براون).

م ت = متوسط الأسئلة المتروكة في آخر الاختبار (وذلك بجمع الأسئلة المتروكة عند جميع الأفراد وتقسيم المجموع على عدد الأفراد).

$ع^2$ خـ = تباين الخطأ (ويحسب برصد عدد الاستجابات الخاطئة عند كل فرد مع إضافة عدد الأسئلة المحذوفة ثم يحسب تباينها بالنسبة لكل الأفراد).

7- معادلة (فلانجان) Flanagan:

$$ر_{11} = \frac{4\ ع_1 \times ع_2 \times ر_{21}}{ع_1^2 + ع_2^2 + 2\ ع_1 ع_2 \times ر_{21}}$$

حيث أن:

ر 11: معامل الثبات للاختبار كله.

ع1 : الانحراف المعياري للجزء الأول من الاختبار.

ع2 : الانحراف المعياري للجزء الثاني من الاختبار.

ر 21: معامل الارتباط بين جزئي الاختبار.

$ع_1^2$: التباين للجزء الأول من الاختبار.

$ع_2^2$: التباين للجزء الثاني من الاختبار.

تستخدم معادلة (فلانجان) في إيجاد معامل الثبات لجزئي الاختبار (للاختبار كله) بعد إيجاد معامل الارتباط بين درجات جزئي الاختبار، وبذلك تستخدم هذه المعادلة للتصحيح كما في معادلة (سبيرمان - براون) لكنها تختلف عنها في العمليات الحسابية.

رابعا: تحليل التباين Variance Analysis:

أن طريقة التجزئة النصفية تصلح لحساب معامل الاتساق الداخلي Coefficient consistency (تجانس نصفي الاختبار) لكنها لا تقيس تجانس الأسئلة ذاتها في الاختبار. إلا أن بعض المختصين يرون أن تجانس الأسئلة لا يدخل في مجال موضوع (ثبات الاختبارات) وإنما هي خاصية من خصائص الاختبارات الجيدة، ومع ذلك لا يمكن استبعاده من مناقشة الثبات، حيث إنه يتناول استقرار استجابات المفحوص على أسئلة أو فقرات الاختبار واحداً بعد الأخر، وعادة ما يسمى هذا النوع من الثبات بـ (معامل الاتساق داخل أسئلة أو فقرات الاختبار) Interitem consistency ويختلف (معامل الاتساق داخل الأسئلة) عن (معامل الاتساق الداخلي) حيث أن معامل الاتساق داخل الأسئلة يشير إلى التجانس الكلي للاختبار بالمعنى المباشر، في حين يكون معامل الاتساق الداخلي للاختبار أقرب إلى معنى التكافؤ كما يتمثل في الصور المتكافئة للاختبار. فقد يكون لدينا اختبارين لهما معاملان للثبات عاليان تم

حسابهما بطريقة الصور المتكافئة أو بالتجزئة النصفية، لكنهما قد يختلفان في درجة تجانس الأسئلة التي يتكون منها صورتا الاختبار أو نصفاه.

وعموماً فإن الفرق بين معامل الاتساق الداخلي - الذي يحسب بالتجزئة النصفية - وبين معامل الاتساق داخل الأسئلة، يمكن اختباره دليلاً على مدى التجانس أو التباين في أسئلة الاختبار.

إن أشهر الطرق لحساب معامل أتساق الأسئلة هي طريقة (كيودر - ريتشاردسون) kuder – Richardson حيث قاما في سنة 1937 بإيجاد معادلة لحساب الثبات هي:

معادلة (KR20) وبالشكل الأتي:

$$ر = \frac{ن}{ن - 1} \left(\frac{مج\ ص\ خ}{ع^2} \right)$$

حيث أن:

ر = معامل الثبات.

ن = عدد فقرات الاختبار.

مج = المجموع بالنسبة لجميع الفقرات.

ص = نسبة الذين أجابوا أجابه صحيحة على الفقرة.

خ = نسبة الذين أجابوا أجابه خاطئة على الفقرة.

$ع^2$ = تباين درجات الاختبار.

وهناك معادلة أخرى لـ (كيودر - ريتشاردسون) وتسمى KR21 وهي:

$$ر = \frac{ن}{ن - 1} \left(\frac{م\ (ن - م)}{ن\ ع^2} \right)$$

حيث أن:

م = الوسط الحسابي لكل الاختبار.

أن المعادلة KR20 قابلة للتطبيق فقط في الاختبارات التي تكون درجة الإجابة على الفقرة أما صحيحة فتأخذ درجة واحدة أو خاطئة فتأخذ صفراً أما المعادلة KR21 فأنها تعطي دائماً تقديراً منخفضاً لمعامل الثبات من أي أسلوب آخر من أساليب التجزئة النصفية، بل وتعطي أكثر انخفاضاً من المعادلة KR20 كما ينبغي مراعاة أن تكون الفقرات متوسطة الصعوبة، أي عدم وجود فقرات صعبة جداً أو فقرات سهلة جداً والا أدى ذلك إلى انخفاض معامل الثبات. إن التجانس الداخلي بين الفقرات يمكن أن تقيسها طريقة (كيودر – ريتشاردسون) وتقوم على تقديرها، أما في حالة وجود مقاييس فرعية فإن هذه الطريقة لا تقيس تجانس جميع الفقرات بل تقدير ثبات المقاييس الفرعية المتجانسة.

تعديل (تيكر) لمعادلة (كيودر - ريتشاردسون):

يهدف تعديل تيكر Tuker إلى تبسيط العمليات الحسابية لمعادلة (كيودر – ريتشارسون) KR20 كما يهدف إلى تجنب عدم الدقة في معادلة (كيودر – ريتشاردسون) KR21 التي تؤدي إلى ثبات منخفض، ووفقاً لتعديل (تيكر) تصبح المعادلة KR20 كالآتي:

$$\text{ر} = \left(\frac{\text{ن}}{\text{ن} - 1}\right)\left(\frac{\text{ع}^2\text{ك} - \text{ن ص}''\,\text{خ}'' + \text{ن ع}^2\text{ ص}}{\text{ع}^2\text{ك}}\right)$$

حيث أن:

ع² ك = تباين الدرجة الكلية على الاختبار.

ص" = متوسط نسبة الإجابات الصحيحة على الفقرة.

خ " = متوسط نسبة الإجابات الخاطئة على الفقرة.

ن = عدد فقرات الاختبار.

$$\text{ع}^2\text{ ص} = \frac{\text{مج ص}^2}{\text{ن}} - \text{ص}'^2$$

حيث أن

ص^2 = مربع نسبة الإجابة الصحيحة على الاختبار

وكما يلاحظ من هذه المعادلة فإن تعديل (تيكر) ينصب على إضافة متوسط نسبة تباين الإجابات الصحيحة على الاختبار في معادلة (كيودر – ريتشاردسون) kr21 مما يؤدي إلى تعديل نتيجتها لتطابق معادلة kr_{20}، ولذلك يفضل استخدام تعديل (تيكر) بدلاً من معادلتي (كيودر – ريتشاردسون) (KR20، KR21).

تعديل (دريسيل) لمعادلتي (كيودر – ريتشاردسون)

بينما نجد أن معادلتي (كيودر – ريتشاردسون)، ومعادلة (تيكر) تقبل التطبيق في الحالات المعتادة بالنسبة للاختبارات المتجانسة والتي تصمم لقياس وظيفة واحدة فقط، لذلك فأنها غير مناسبة للاختبارات التي تتضمن نظاماً للتصحيح يقوم على أوزان متدرجة من الفقرات. حيث يمكن استخدام معادلتي (كيودر – ريتشاردسون) في حالة الإجابة على الفقرة أما مقبولة أو مرفوضة، صواب أو خطأ، نعم أو لا، أما إذا كانت الإجابة متدرجة على أوزان معينة فإن المعادلة المناسبة هنا لحساب الثبات بنفس منطق معادلتي (كيودر – ريتشاردسون) أو ثبات الاتساق الداخلي وتجانس الفقرات هي معادلة (دريسيل) والتي نصها:

$$\text{ر}_{\text{ن}} = \left(\frac{\text{ن}}{\text{ن} - 1}\right)\left(\frac{\text{ع}^2\text{ك} - \text{مجـ و}^2_{\text{ق}}\ \text{ص}_{\text{ق}}\ \text{خَ}_{\text{ق}}}{\text{ع}^2\text{ك}}\right)$$

حيث أن:

و ق = الوزن الخاص بالاستجابات الصحيحة للفقرة (ق).

ص ق = نسبة الإجابات الصحيحة للفقرة (ق).

خ ق = نسبة الإجابات الخاطئة للفقرة (ق).

وعموماً فإن معادلتي (كيودر – ريتشاردسون) أو تعديلاتها لا تصلح لاختبارات السرعة الموقوتة وبذلك تكون مقتصرة على اختبارات القوة فقط.

8- معادلة (كرونباخ العامة) للثبات:

يقدم (كرونباخ) Cronbach معادلة عامة تعد المعادلات السابقة جزءاً منها، أو حالة خاصة منها، وهي تنطلق من المنطق العام لثبات الاختبار ويطلق على معادلة (كرونباخ) اسم معامل (ألفا) Alpha، والتي صيغتها:

$$\alpha = \left(\frac{\text{ن}}{\text{ن} - 1}\right)\left(1 - \frac{\text{م ع ق}}{\text{ع ك}}\right)$$

حيث أن:

ع ق = تباين الجزء ق من الاختبار – مهما كان طول هذا الجزء.

ع ك = التباين الكلي للاختبار.

ن = عدد أجزاء الاختبار.

وتنطبق هذه الصيغة لحساب الثبات سواء أكانت أجزاء الاختبار عبارة عن نصفين أو تتعدد إلى أن تكون فقراته جميعها كما في معادلة (كيودر – ريتشاردسون). ويفضل استخدام هذا المعامل عندما يكون الهدف تقدير معامل ثبات المقاييس الجوانب الوجدانية والشخصية نظرا لأنها تشتمل على مقاييس متدرجة لا يوجد بها اجابة صحيحة وأخرى خاطئة.

ويجب إلا يفوتنا أن نذكر بأن هناك طرقاً أخرى لقياس معامل الثبات في حالة ما إذا كانت الإجابات متعددة وليست ثنائية ولعل أهمها وأشهرها الطرق التي تعتمد على تحليل التباين والتحليل العاملي.

خامسا: ثبات المصححين Scorer Reliability:

من المصادر الرئيسة لتباين الخطأ في درجة الاختبار أو المقياس الذي يعتمد

على تقييم المصحح وليس على مفتاح للتصحيح أو مجرد عند الاستجابات، ما يطلق عليه عدم ثبات التقييم أو التصحيح، وهذا ما يحدث عادة في الاختبارات الإسقاطية التي تعتمد أساساً على تأويل استجابة المفحوص للمنبه وفقاً لمحاكاة معينة، لكن انطباقه لتأويل على شروط المحاكات كثيراً ما يرجع إلى ذاتية المصحح رغم التزامه بشروط المحك، مما تظهر الفروق بوضوح عندما يستخدم مصححان أو أكثر نفس المحك لتصحيح نفس العينة من الاختبارات، كما تظهر نفس الحالة أيضاً في الاختبارات الخاصة بقياس القدرات الإبداعية، حيث أن المصحح هو الذي يحدد مدى جودة أو أصالة الاستجابة في اختبارات الأصالة رغم توفر معايير التصحيح.

كما أن الاختبارات التحصيلية لاسيما في اختبار المقال (essay) يمكن تقدير درجتها على الحكم الذاتي للمصحح مما يمكن أن تصحح إجابات المفحوصين بواسطة أثنين أو أكثر من الحكام أو الخبراء، ومن ثم إيجاد معامل ثبات المصححين على نفس المجموعة. بيد إنه نادراً ما تحتاج الاختبارات الموضوعية إلى حساب معامل ثبات المصححين (Scorer Reliability) وذلك يعود لقلة أخطاء التصحيح فيها ولسهولتها في إعطاء الدرجة وتقدير عدد وحجم الأخطاء. إن معامل ثبات المصححين في هذه الحالة هو (معامل الارتباط) بين مجموعتي الدرجات التي يحصل عليها المفحوصين في العينة التي جاءت من تصحيح الباحثين المستقلين أو المصححين كل على انفراد، إلا أن إيجاد معامل ثبات المصححين لا يعني عدم الحاجة إلى معامل ثبات المجيب أو معامل ثبات الاختبار، بل ينبغي إيجاد معامل ثبات المجيب والاختبار إضافة إلى معامل ثبات المصححين، حيث أن معامل ثبات المصححين يشير إلى أي مدى يتفق اثنان أو أكثر في تقدير درجات نفس المجموعة على نفس الاختبار، بينما يشير معامل ثبات الممتحن (المجيب) إلى مدى تناسق أداء المجيبين على نفس مجموعة الأسئلة أو الفقرات، في حين يشير معامل ثبات الاختبار إلى مدى تشابه أداء الممتحنين على الفقرات أو الأسئلة المختلفة، ويفضل استخدام أسلوب Test – Retest في إيجاد معامل ثبات الممتحنين (المجيبين).

الفصل السادس

المعاييـــر

يحتوي الفصل على:

- الحاجة إلى المعايير
- خصائص المعايير
- المعايير والتقنين
- أغراض المعايير
- أنواع المعايير
- الدرجة المعيارية

مقدمة:

أن المشكلة التي تواجهنا في قياس السلوك البشري هي في البحث عن وحدات ذات دلالة ومعنى يمكن تفسيرها ومقارنتها، والمشكلة أيضاً هي في وضع مستويات بها نقارن ونقيس ونقوم ونفسر ونعلل السلوك الإنساني، هذه الوحدات هي المعايير.

الحاجة إلى المعايير:

أن الدرجات الخام التي يحصل عليها الطلاب في اختبار ما وكذلك نسبة الطلبة الناجحين في الاختبار إلى المجموع الكلي ليس لها في ذاتها معنى أو دلالة. فهي لا

تحدد مركز الشخص بالنسبة لمجموعته أو بالنسبة لعينة التقنين، ولهذا لا بد من الرجوع إلى معيار يحدد معنى هذه الدرجة أو هذه النسبة المئوية. وهكذا فإن هذا المعيار سيدلنا على مركز الشخص بالنسبة للمجموعة، هل هو متوسط، فوق المتوسط، أقل من المتوسط وما مدى هذا البعد.

خصائص المعايير:

1- أن يكون للدرجة الواحدة معنى موحد من اختبار لاختبار.

2- أن تكون هذه الوحدات متساوية بحيث أن (10) درجات على جزء من اختبار تدل على نفس الشيء الذي تعنيه (10) درجات على جزء آخر من نفس الاختبار.

المعايير والتقنين:

يشتق المعيار من عينة التقنين التي تمثل المجتمع الأصلي المدروس. فتكون الدرجات المحولة للدرجات الخام لأفراد هذه العينة مصدراً للمعيار. أي أننا عندما نقارن درجة المفحوص في اختبار ما نعود إلى من يكافئه في عينة التقنين.

وبناءً على ذلك فإن المعايير هي نتائج إجراء الاختبار على عينات التقنين. وعند نقل الاختبار من بيئة إلى أخرى يجب تقنينه على عينات موازية لعينات التقنين الأصلية لاشتقاق المعايير الجديدة.

أغراض المعايير:

1- تحدد مركز الفرد بالنسبة لعينة التقنين.

2- يمكن من خلالها مقارنة الفرد على مقياس بمركزه على مقياس آخر.

وهناك طريقتان نربط بهما درجة الفرد بإطار عام وهي:

1- المقارنة بسلسلة متدرجة من الجماعات ونبحث عن أتباعه التي تكافؤه عن طريق متوسطها وكل جماعة تمثل أما صفاً دراسياً معيناً أو عمراً زمنياً معيناً.

2- نحدد موضع الفرد في جماعة معينة بالإشارة إلى النسبة المئوية من الجماعة التي يمتاز بها أو بالرجوع إلى متوسط الجماعة وانحرافها المعياري.

أنواع المعايير:

المعايير القومية (الوطنية):

وهي أكثر أنواع المعايير استخداماً وهي تخص التربية ومستوى العمر. ومن المشاكل التي تواجهها هي كثرة المتغيرات التي يجب أخذها بنظر الاعتبار في تحديد المعيار، وتحسب عن طريق جمع البيانات عن البلد عموماً.

المعايير المحلية:

وتحسب عن طريق جمع البيانات من منطقة تعليمية معينة واستخدامها بدلاً عن المعايير الوطنية في تقدير أداء الفرد.

معايير العمر:

وهي قيم تمثل الأداء الفعلي أو المتوسط لأشخاص من أي مستوى عمر زمني معين في أي صفة مقاسة. ويمكننا أن نعد معياراً عمرياً لأي سمة تنمو مع زيادة السن. فمعيار العمر هو القيمة المتوسطة لهذه الصفة لأشخاص من نفس العمر.

ومن عيوب هذه المعايير أن بعض الصفات لا تتناسب فيها الزيادة كالزيادة في العمر كالوزن والذكاء والطول. كما أنها لا تفيد في تقدير سمات الأفراد المتطرفين زيادة أو نقصاناً. كما لا يمكن مقارنة منحنيات النمو للسمات المختلفة مع بعضها لاختلاف سرعة النمو من صفة لأخرى لدى الفرد. ولتلافي ذلك تم استخراج نسبة الذكاء، ونسبة التعليم، ونسبة التحصيل.

معايير الصف:

وهي تطوير لمعايير العمر وهي متوسط درجات أفراد الصف الواحد في اختبار معين. وهي أسهل لأنها مبنية على مجموعات موجودة أصلاً في النظام المدرسي.

حساب معايير العمر والصف:

1- يعطى الاختبار المراد عمل معايير عمر أو معايير صف له ممثلة فيه الأعمار المختلفة بالنسبة لمعايير العمر وممثلة فيه الصفوف المختلفة بالنسبة لمعايير الصفوف.

2- يحسب الوسيط أو المتوسط التحصيلي لكل عمر أو لكل صف على أساس أن الفرق يساوي سنة واحدة لكل منهما.

3- يرسم منحنى بياني للمتوسط أو الوسيط بحيث تمثل الأعمار أو الصفوف على المحور السيني والمتوسطات على المحور الصادي.

4- يرسم منحنى يمر بمواضع أقرب ما يمكن من مواضع النقاط الممثلة للمتوسطات أو الوسيطات.

5- يمر طرفا المنحنى عند نهايتيه لكي يمكن تفسير الدرجات التي تفوق هذه المتوسطات أو تقل عنها.

إن استخدام معايير العمر والصف يتطلب وجود اختبارات مقننة بدقة، كما يجب توفر عدداً كافياً منها لدى المعلم، وفي كثير من الأحيان لا تتوفر هذه الاختبارات، كما أنها لا تستخدم إلا في دراسة مستويات السمات والنواحي التي تتغير بوضوح مع العمر، لذا فمن الضروري استخدام طرق أخرى لتحويل الدرجات الخام إلى درجات يمكن مقارنة بعضها بالبعض الأخر.

ويمكن تصنيف هذه الطرق إلى ثلاث مجموعات رئيسة:

1- طرق مبنية على حساب ترتيب التلميذ بالنسبة لمجموعته في الصف أو بالنسبة لعينة التقنين كالمئينات والأعشاريات.

2- طرق مبنية على التحويل الخطي للدرجات الخام، كالدرجات المعيارية.
3- طرق مبنية على التحويل المساحي للدرجات الخام، كالدرجة التائية، والمعيار الجيمي، والتساعي المعياري والسباعي المعياري.

المئينات (المعايير المئينية):

في المعايير المئينية نقارن الفرد بالآخرين من عمره أو صفه لتحديد مركزه بالنسبة لجماعته، فنعرف مثلاً أن الفرد يفوق 75 ٪ من جماعته في مادة معينة.

فالدرجة المئينية هي الدرجة التي تقع تحتها نفس النسبة من الأفراد، وهي النقط التي تقسم التوزيع التكراري إلى أجزاء متساوية، ويعتبر الوسيط هو النقطة التي تقسم التوزيع التكراري إلى قسمين متساويين أي هو المئيني الخمسين.

وفي المئينيات نحول البيانات إلى السلم المئوي أي توزع البيانات بحيث تشمل الفترة بين الصفر والمائة. أي أن المئيني صفر يمثل الحد الأدنى للفئة التي توجد عند بداية التوزيع بينما المئيني (100) يمثل الحد الأعلى للفئة التكرارية التي تقع عند نهايته ويجب أن نفرق ونميز بين المئيني وبين الرتبة المئينية والنسبة المئوية. فالمئيني هو الدرجة المقابلة لمرتبة مئينية معينة، أما النسبة المئوية فهي تشير إلى النسب المئوية لدرجات الاختبار وليس الأفراد.

وتحسب المئينيات وفقاً للمعادلة التالية:

$$\text{المئيني} = ح_1 + \frac{\frac{م\ ن}{100} - ت\ ج\ ق}{ت} \times ف$$

حيث أن:

$ح_1$ = الحد الأدنى للفئة التي بها المئيني المطلوب.

م ن = ترتيب المئيني × مجموع التكرار.

ت ج ق = التكرار المتجمع للفئة السابقة لفئة المئيني.

ت = تكرار فئة المثيني.
ف = طول الفئة. م ن

$$\text{ولاستخراج رتبة المثيني ت} = \frac{\text{م ن}}{100}$$

وتتميز المثينيات بأنها تعطي صورة صادقة لترتيب الفرد بالنسبة للمجموعة وأنها سهلة في حسابها وواضحة في مدلولها.

ومن عيوبها أن الوحدات المثينية على منحنى التوزيع غير متساوية إذ تقل المسافات بين المثينيات في الوسط وتزيد كلما اتجهنا بعيداً عن الوسط. فالمسافة بين المثين (صفر) والمثين (10) تساوي سبعة أمثال المسافة بين المثين (40) والمثين (50).

وهذا يعني إنه ليس هناك أتفاق بين توزيع الدرجات الخام وتوزيع المثينيات فكل ما تعطيه المثينيات هو ترتيب الدرجات فقط.

الدرجة المعيارية:

وهي المسافة التي تبعد فيه الدرجة الخام عن الوسط الحسابي معبراً عنها في وحدات من الانحراف المعياري.

وتسهل الدرجة المعيارية مقارنة المجموعات مع بعضها حتى وأن اختلفت متوسطاتها.

$$\text{الدرجة المعيارية} = \frac{\text{الدرجة الخام - المتوسط الحسابي}}{\text{الانحراف المعياري}}$$

ومن هنا يتضح أنها تستخدم الانحراف المعياري في حين أن المعايير السابقة لم تستخدمه وغالباً ما تقع درجات التوزيع بين (+3) و (-3) درجة معيارية، فالدرجة الخام التي تكون أعلى من المتوسط تكون إشارتها موجبة والأقل من المتوسط تكون أشارتها سالبة.

ورغم فوائد الدرجة المعيارية فهي لا تصلح لعملية المقارنة إلا إذا كان التوزيع اعتدالياً أو قريباً منه ونظراً لكثرة درجاتها السالبة فقد عمد إلى تعديل الدرجة المعيارية للتخلص من الإشارات السالبة وتسوية المنحنى، حيث تكون جميع الدرجات المعيارية موجبة. باستخدام الدرجات المعيارية المعدلة مثل:

أ - الدرجة التائية:

حيث تحول الدرجات المعيارية إلى درجات معيارية متوسطها (50) وانحرافها (10) وقد عدلت هذه الدرجة أيضاً بمضاعفة الوسط الحسابي والانحراف المعياري بحيث أصبحا (20)، (50) و (100) على التوالي.

ب ـ الدرجة الجيمية:

وهدفها إيجاد درجات معيارية تناسب بعض الاختبارات التي تقيس وظائف تكون الفروق بين الأفراد فيها واضحة. وتعدل فيها الدرجات المعيارية بحيث يكون المتوسط (5) والانحراف المعياري (2)، وقسم التوزيع فيها (11) قسماً.

فإذا كانت الدرجة المعيارية لفرد هي 0,4 فتصبح الدرجة الجيمية له

.(5,8 = 5 + 2 × 0,4)

جـ التساعي المعياري:

ويقوم على أساس الدرجة الجيمية حيث أن مستوياته أقل من تقسيم الدرجة الجيمية فبدلاً من أن يكون (11) قسماً قسم إلى (9) أقسام إذ إنه يجمع بين الدرجة الجيمية المساوية للصفر والدرجة الجيمية المساوية لواحد صحيح، كما يجمع بين الدرجة الجيمية (9) والدرجة (10) في درجة واحدة هي (9) ويوضح هذا المعيار إذا كان الأفراد المتطرفون متقاربين سلباً أو إيجابياً. والجدول التالي يوضح المقارنة بين المعايير.

نوع المعيار	نوع المقارنة	نوع الجماعة
1- معايير العمر	مقارنة الفرد بالجماعة التي يكافئها	جماعات متتابعة في العمر.
2- معايير الصف	مقارنة الفرد بالجماعة التي يكافئها	جماعات متتابعة في الصفوف
3- معايير المئينات	النسبة المئوية من الجماعة التي يمتاز عنها الفرد.	جماعات من العمر أو الصف التي ينتمي إليها الفرد.
4- معايير الدرجة المعيارية	عدد الانحرافات المعيارية التي يزيد أو يقل بها الفرد عن متوسط الجماعة.	جماعات من نفس العمر أو الصف التي ينتمي إليها الفرد

وهناك طرق منها:

أ. نسبة الذكاء I. Q. :

وحسابها بالمعادلة الآتية:

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلي}}{\text{العمر الزمني}} \times 100$$

ب- النسبة التعليمية E. Q. :

وحسابها:

$$\text{النسبة التعليمية} = \frac{\text{العمر التحصيلي}}{\text{العمر الزمني}} \times 100$$

ويحسب العمر التحصيلي بعد معرفة متوسط كل صف دراسي والدرجة التي يحصل عليها الفرد تحول إلى عمر تحصيلي. فالطالب (س) الذي يحصل على الدرجة(8) وعمره الزمني (10) سنوات نرجع إلى متوسط الصفوف ولنفرض أن متوسط الصف الرابع (6) ومتوسط الصف الخامس (10) فلحساب العمر التحصيلي نستخرج أولاً مدى السنة الدراسية بين الرابع والخامس:

(10 – 6 = 4) ثم نستخرج المدى الذي قطعه الطالب من السنة الدراسية (8 – 6 = 2) فيكون الصف فيه الطالب (س) هو:

$$= 4 + \frac{6-8}{6-10}$$

$$= 4 + \frac{1}{2} = 4{,}5$$

أما النسبة التعليمية فهي $= \frac{4{,}5}{10} \times 100 = 45$

ج- النسبة التحصيلية A. Q. وحسابها:

نسبة التحصيل = $\frac{\text{العمر التحصيلي}}{\text{العمر العقلي}} \times 100$

أو = $\frac{\text{النسبة التعليمية}}{\text{نسبة الذكاء}} \times 100$

فإذا كان العمر التحصيلي لفرد هو(12) سنة وعمره العقلي (10) سنوات فإن نسبة التحصيل

$$= \frac{12}{10} \times 100$$

فتفسر هذه الدرجة ونقول أن هذا الطالب حصل أكثر مما يسمح به مستوى مجموعته العقلي وعلى هذا فنخمن إنه يبذل مجهوداً أكثر من الطالب المتوسط.

وإذا كانت النسبة التعليمية لفرد (100) ونسبة ذكائه (125) فالنسبة التحصيلية

$$= \frac{100}{125} \times 100 = 80$$

وهناك معايير أخرى منها:

أ- معايير الأداء النوعية: يتعرف من خلالها المعلم على المادة التي تحتاج إلى تأكيد أكبر.

ب- معايير المتوسط المدرسي: وتستخدم لمقارنة متوسط المدرسة بالمدارس الأخرى المجاورة.

ج- التقديرات النوعية: وتستخدم تقديرات معينة وليس الدرجات على أساس أن هناك مدى من الدرجات لكل تقدير مثال ذلك: (ممتاز، جيد، وسط، ضعيف).

د- معايير الدراسة الخاصة: وتستخدم للمقارنة بين أقسام متشابهة أو مرحلة صفية/ الخ حيث يكون متوسط مجموع الأقسام أو الكليات ذات الاختصاص الواحد هو المعيار وعلى سبيل المثال متوسط جميع طالبات كليات التمريض، متوسط جميع طلاب الهندسة .... الخ.

وبشكل عام فهناك نوعان معياريان، أو نوعان من المجموعات المرجعية للمقارنات:

1- استخدام مجموعة مرجعية مفردة ويتم وصف درجة الفرد التي حصل عليها من خلال علاقتها مع ترتيب درجات تلك المجموعة.

2- استخدام درجات الوسط الحسابي لمجموعات مرجعية متسلسلة ويتم وصف درجة الفرد التي حصل عليها من خلال تطابقها أو اقترابها من متوسط إحدى المجموعات.

الفصل السابع

الاختبارات المحكية المرجع والاختبارات المعيارية المرجع والاختبارات التشخيصية

يحتوي الفصل على:

- مقدمة
- الاختبارات محكية المرجع والاختبارات المعيارية المرجع
- أوجه الاختلاف بين الاختبارات معيارية المرجع ومحكية المرجع
- تصنيفات الاختبارات محكية المرجع Categories of CRT
- خطوات بناء الاختبارات محكية المرجع (الجانب المعرفي)
- طرق تقدير صدق الاختبارات محكية المرجع
- بناء الاختبارات محكية المرجع لقياس الجانب الوجداني
- طرق تقدير ثبات الاختبارات محكية المرجع
- مسميات وتعريف درجة القطع Definitions of cut off score
- طرق تحديد درجة القطع
- أسس تقويم طرق تحديد درجة القطع
- الاختبارات التشخيصية
- خطوات بناء الاختبارات التشخيصية
- أنماط التعليم ومستوياته

مقدمة

وجه العديد من علماء القياس التربوي انتقاداتهم إلى الاختبارات المعيارية المرجع لكون أن معظم هذه الاختبارات نشأت مرتبطة بالفلسفة التربوية التي كانت سائدة في الولايات المتحدة الأمريكية في أوائل القرن العشرين، وهي تصنيف الأفراد بحسب مركزهم النسبي بين أقرانهم في القدرات المختلفة، وبخاصة عندما دعت الحاجة إلى تصنيف الضباط والجنود قبل اشتعال الحرب العالمية الأولى، حيث ظهر في ذلك الوقت اختبار آرثر أوتيس A. Otis وهو أول اختبار جمعي، وقد أدى تحقيق هذا الاختبار للغرض الذي وضع من اجله إلى تشجيع حركة الاختبارات جماعية المرجع لتصنيف الأفراد تبعاً لقدراتهم وتحصيلهم وغير ذلك من الصفات. كما استخدمت في مجالات التوجيه المهني، والتشخيص الإكلينيكي، وانتقاء الأفراد للوظائف المختلفة وتحديد الفرق في تحقيق كثير من الأغراض، إلا إنه تبين قصورها في تقويم فاعلية المنظومة التعليمية، وفي قياس المهارات المختلفة المطلوب تحققها لدى الدراسية المناسبة للطلاب، وقياس فاعلية البرامج التربوية. وعلى الرغم من نجاح هذه الاختبارات الأفراد.

كما بينت الدراسات المستفيضة حول الاختبارات المقننة المستخدمة في قياس الذكاء أن كثيرا منها يشتمل على مفردات تتحيز ضد مجموعة أقلية تعيش في نفس المجتمع الذي استمدت منه معايير هذه الاختبارات. كما أنها تؤدي في كثير من الأحيان إلى التفرقة العنصرية بين الأجناس. وقد ترتب على ذلك العديد من المشكلات الاجتماعية والتربوية مثل التشخيص غير الدقيق للأطفال المتخلفين عقلياً والاختيار غير الصائب لطلاب الدراسات العليا أو المتقدمين للحصول على المنح الدراسية.

كما أن كثير من المقاييس المقننة للشخصية المستخدمة في انتقاء أفضل المتقدمين لشغل وظائف معينة لا تتميز بالثبات والصدق اللازم توفرهما بقدر كبير من الثقة. وكثير من المفردات والعبارات التي تشتمل عليها وبخاصة تلك التي تتعلق بالأمور

الجنسية والسياسية والدينية تمس الحرية الفردية للأشخاص الذين تطبق عليهم هذه المقاييس.

وأن الاختبارات التحصيلية المقننة لا تقيس الأهداف التعليمية التي يسعى المعلم إلى تحقيقها في الفصل المدرسي. كما أنها لا تقدم الفائدة المرجوة في تشخيص مواطن الضعف والقوة في تحصيل الطلاب حتى يستطيع المعلم أن ييسر وسائل المعالجة ورفع كفاءة العملية التعليمية.

وقد استثارت أوجه النقد هذه المربين وعلماء القياس إلى أهمية استخدام مدخل آخر لقياس نواتج التعلم أطلق عليها "القياس محكي المرجع Referenced Criterion Measurement". وفرق جليزر بين نوعين من القياس أحدهما لقياس المعياري المرجع Norm Referenced Measurement، ويتركز الاهتمام فيه على ترتيب الفرد بالنسبة لأقرانه في القدرة التي يقيسها الاختبار، والآخر هو القياس المحكي المرجع الذي يهدف إلى موازنة أداء الفرد بمستوى أداء مطلق أو مستوى كفاية معين دون الحاجة إلى موازنة أدائه بأداء الأفراد الآخرين.

وقد أثار جليزر كثيرا من المناقشات والجدل بين علماء القياس بعامة والمتخصصين في تطبيقات تكنولوجية التعليم بخاصة، ولكن لم يحدث نشاط علمي ملحوظ نحو تحقيق ما نادى به جليزر إلا حوالي عام 1969.

ففي عام 1969 بدأ جيمس بابام James Popham عالم النفس الأمريكي بجامعة كاليفورنيا يتزعم حركة القياس محكي المرجع. ودعا إلى عقد مؤتمر متخصص في أمريكا عام 1970 لمناقشة القضايا والمشكلات السيكومترية المتعلقة بهذا المفهوم من القياس. ونشر له اول كتيب عن مقدمة في القياس المحكي المرجع عام 1971. وترتب على ذلك حركة بحثية كبيرة منذ ذلك الوقت ولحد الآن. فقد صدر كتاب لنورمان جرونلند N. Grounlund عن إعداد الاختبارات المحكية المرجع والكتاب موجه للمعلم المهتم ببناء هذا النوع من الاختبارات في المستويات المعرفية المختلفة. كما ويعتبر كتاب جيمس بابام Popham الذي صدر عام 1978 أول كتاب متكامل

عن القياس محكي المرجع وتلى ذلك صدور عدة كتب من أهمها كتاب بيرك Berk عام 1980 وآخر في عام 1983.

الاختبارات محكية المرجع والاختبارات المعيارية المرجع

تعتمد عمليتي القياس والتقويم التربوي على أدوات ومقاييس تقدم بيانات كمية تتيح للتربويين فهم الظاهرة التربوية بشكل دقيق ومحدد، وتعتبر الاختبارات والمقاييس التربوية جزءً هاماً من أجزاء العملية التربوية فمن خلالها يستطيع المعلم الحكم على مدى تحقق أهداف البرامج التعليمية والتدريبية التي يقوم بتدريسها، كما أنها تفيد الطالب أيضاً في التعرف على مستوى تحصيله الدراسي من خلال أدائه في الاختبار. كما تتعدد استخدامات الاختبارات في المؤسسات التربوية فمنها ما يستخدم للقياس والتقويم أو التشخيص ...الخ.

وبنظرة سريعة على حركة بناء المقاييس التربوية المعاصرة يلاحظ أن هناك ثلاثة اتجاهات رئيسة يختلف كل اتجاه منها عن الآخر في الافتراضات التي يقوم عليها، وفي الإطار المرجعي الذي يعتمد عليه تفسير نتائج الطلاب في الاختبار، وفي إعداد أدوات القياس وهي القياس معياري المرجع Norm Referenced Measurement، والقياس محكي المرجع Criterion Referenced Measurement، ونظرية السمات الكامنة Latent Traits Theory.

وأشار كل من جيلزر وبابام وهيوزيك (Popham & Husek, 1969; Glaser, 1994) بأن الاتجاه الأول يعتمد على تفسير درجات الطالب في ضوء معايير محددة (Norms). أي ترد درجات الطالب في الاختبار إلى جداول معيارية ومن ثم يحدد مستوى أداء الطالب النسبي بناءً على موقع درجة الطالب مقارنةً بمتوسط تحصيل المجموعة المعيارية في الاختبار، ويركز هذا الاتجاه على الفروق الفردية بين الطلاب بحيث يوضح مدى الفارق بين مستوى طالب وآخر، أي إنه يقيس مقدرة طالب معيّن مقارنةً بمقدرة طلبة آخرين.

وتعرض الاتجاه معياري المرجع لبعض أوجه النقد أشير إليها في أدبيات القياس التربوي ومنها ما ذكرها بابام وهيوزيك Popham & Husek 1969 حيث تبين أن الاختبارات معيارية المرجع تعتمد على مقارنة أداء الطالب بأداء مجموعته الصفية، وبالتالي فإنه من المحتمل أن يختلف موقع الطالب باختلاف خصائص مجموعته المعيارية، بالإضافة إلى إنه ليس بالضرورة أن يمثل متوسط المجموعة المعيارية الأداء المطلوب للنجاح، وتهتم هذه الاختبارات أيضاً بإبراز الفروق الفردية بين الطلاب دون النظر إلى مدى إتقان الطالب للمهارات والمعلومات المراد قياسها وذلك لتأهيله لبرامج تدريبية وتعليمية جديدة، وبالتالي فهي لا تساعد على تقويم وتشخيص كل من كفاية الطالب وبعض البرامج التدريبية (برامج التعليم الفردي - البرامج القائمة على الكفايات بالقدر المطلوب). أي أنها لا تزود المعلم بمعلومات دقيقة تساعد على اتخاذ القرارات التربوية المناسبة حول مستوى كفاية الطالب والمنهج الدراسي.

كما أكد بابام (Popham 1978) بأن هذه الانتقادات لعبت دوراً في ظهور تطورات جديدة في منهجية الاختبارات والمقاييس التربوية أدت إلى ظهور اتجاه آخر أطلق عليه القياس محكي المرجع Criterion-Referenced-Measurement، حيث تسجل أول مقالة لجليزر في عام 1963م تناقش هذا الاتجاه.

كما أشار كل من هامبلتون وآخرون (Hambleton, et al 1976) إلى أن هذا الاتجاه لا يعتمد على تحديد مستوى الطالب النسبي داخل مجموعته الصفية ولكنه يعتمد على تحديد ما يستطيع وما لا يستطيع الطالب أن ينجزه. أي ينسب أداء الطالب إلى محتوى الاختبار ذاته، ويعتمد هذا المقياس في بنائه على تحديد نطاق الأهداف السلوكية تحديداً دقيقا، ومن ثم قياس مدى إتقان الطالب لهذه الأهداف من خلال مقارنة أداء الطالب بمستوى معين من الأداء (درجة القطع).

وأضاف ايبل وهامبلتون وايجنور (Hambleton & Eignor, 1979; Ebel, 1979) إلى أن القياس محكي المرجع يهدف إلى تصنيف الطلاب إلى مجموعتين الأولى متقنة

masrey والثانية غير متقنة Non-Mastery للمهارات والمعارف وذلك اعتماداً على (مستوى أداء - معيار - محك - درجة القطع - درجة النجاح) محددة تحديداً مسبقاً. ويتمثل هذا المحك (Criterion) في نسبة عدد الأسئلة التي يجب أن يقوم الطالب المتقن بالإجابة عليها بصورة صحيحة، وبالتالي إذا كانت درجة الطالب في الاختبار تزيد أو تساوي درجة القطع (Cut off Score) اعتبر الطالب متقناً للمهارات وإذا لم تكن كذلك اعتبر الطالب غير متقن. يتضح من ذلك أن عملية تصنيف الطلاب (Classification) إلى فئتين تعتمد على قيم درجات قطع الاختبارات، لذا يجب تحديد هذه الدرجة بعناية فائقة جداً لأنه يترتب على تحديدها اتخاذ قرارات هامة منها: تقويم مستوى تحصيل الطالب في الاختبار، وتقويم البرامج والمناهج التعليمية والتدريبية ومن ثم تشخيصها بصورة مناسبة

ويعتبر تحديد درجة القطع (النجاح) أمراً بالغ في الصعوبة و الأهمية لذلك كثرت المداخلات بين علماء القياس التربوي حول كيفية تحديد درجة القطع المثلى التي يمكن على ضوئها اتخاذ القرارات المناسبة، نتج عن تلك المداخلات طرائق متعددة ومختلفة لتحديد درجة القطع في الاختبارات محكية المرجع، والتي ذكر بيرك (Berk, 1986) أنها بلغت ثمان وثلاثين طريقة وتقسم طرق تحديد درجة القطع إلى ثلاثة أقسام:

1- يحتوي القسم الأول على الطرق التحكيمية Judgmental - Methods حيث تعتمد هذه الطرق على آراء) المحكمين) المختصين في المجال أو المحتوى المراد قياسه لتحديد درجة قطع الاختبار.

2- يحتوي القسم الثاني على طرق تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية Methods–Empirical – Judgmental وهي تعتمد على آراء المحكمين مع الاعتماد على بيانات ترتبط بالأداء الفعلي للطلاب.

3- ويحتوي القسم الثالث على طرق تعتمد بشكل رئيس على البيانات التجريبية

وجزئياً على التحكيم حيث تعتمد هذه الطرق على أداء الطلاب الفعلي في الاختبارات والتحليلات الإحصائية، وينحصر دور عملية التحكيم في انتقاء الطلاب المتقنين وغير المتقنين من الطلاب.

وأشار أبو علام (1995) إلى أن كثرة الانتقادات الموجهة إلى القياس معياري المرجع كانت سبباً في ظهور نظرية السمات الكامنة (LTT)، ومن تلك الانتقادات ما يلي:

اعتماد بناء جميع خصائص الاختبارات السيكومترية على خصائص عينة الدراسة، وعلى صعوبة عينة مفردات الاختبار، يفترض خطية ميزان درجات الطالب في الاختبار غير إنه عادة ما يكون الميزان منحنياً لأن الفرق الثابت بين درجتين من درجات الاختبار يختلف تفسيره بناء على موقع الدرجات على متصل السمة أو القدرة المراد قياسها، كما يفترض أن العلاقة بين زيادة درجة الطالب في الاختبار وبين مقدار السمة أو القدرة علاقة خطية مطردة. وقد لا يكون هذا الافتراض صائباً، إذا لم يحصل بعض الطلاب من ذوي القدرات المتميزة على درجات مرتفعة في الاختبار وفي المقابل يمكن أن يحقق بعض الطلاب من ذوي القدرات المتدنية على درجات مرتفعة في الاختبار، كما تتأثر مفردات الاختبار بالظروف المحيطة به حيث يمكن أن تتغير درجات الطلاب بحذف أو تغيير أي مفردة من مفردات الاختبار الأمر الذي يؤثر في تمثيل المفردات للنطاق السلوكي المراد قياسه.

كما أن هذه النظرية وفرت نماذج تقدر السمات الكامنة (القدرة) في الفرد بصورة مستقلة عن أفراد العينة الدراسية.

وتوجد ثلاثة أبعاد توفر الأسس الموضوعية في نماذج السمات وهي: بعد السمة المقاسة الذي يعرف بواسطة مجموعة من البنود وهي أما أن تكون أحادية (صعوبة المفردة) أو ثنائية (صعوبة المفردة ومعامل التمييز) أو ثلاثية المعلم (صعوبة

المفردة ومعامل التمييز وعامل تخمين الإجابة)، وبعد استقلالية القياس حيث لا يعتمد تقدير صعوبة البند على صعوبة البنود الأخرى، كما لا يعتمد أيضاً تقدير قدرة الأفراد على قدرة أي مجموعة أخرى.

وأما توازي المنحنيات المميزة للبنود الذي يوفر الحد الذي تميز فيه المفردات بين الأفراد ذوي المستويات المختلفة في قدرة ما.. و أشار علام (1995) إلى نماذج النظرية الثلاثة وهي نموذج (راش) ويعد نموذجاً أحادي المعلم، ونموذج (لورد) ويعد ثنائي المعلم ونموذج (بيرنبوم) ويعد ثلاثي المعلم.

وكما أشرنا سابقا إلى أن الاختبارات التحصيلية تهدف إلى:

- التأكد من توافر الحد الأدنى اللازم من مستويات الأداء بالنسبة لعمل أو مهنة معينة.
- التصنيف المهني والتعليمي والعسكري.
- مقارنة أداء الطالب بأداء أقرانه.
- تحفيز دافعية التلاميذ على التحصيل.

ولكي تحقق الاختبارات التحصيلية الأهداف المنشودة، يجب أن يكون الاختبار معداً بطريقة جيدة بحيث يغطي عينة ممثلة من النواتج التعليمية والمادة الدراسية، وأن يصاغ بطريقة تتناسب مع الغايات التي ستستخدم النتائج من أجلها وأن يتمتع بقدر كبير من الصدق والثبات، وأن يتوخى الحذر في تفسير نتائجه وذلك بسبب أخطاء القياس، وأن يستخدم لتحسين مستوى تعلم الطلاب وأن يحتوى على أسئلة ذات نوعية ملائمة لقياس الناتج التعليمي المطلوب.

وتوجد العديد من تصنيفات الاختبارات التحصيلية وذلك حسب نوع الاستجابة وطريقة التطبيق وصياغة الأسئلة وتفسير الدرجات الناتجة، ولعل أهم تصنيف للاختبارات التحصيلية من حيث طريقة تفسير الدرجات الناتجة من تطبيق الاختبار كالتالي:

الاختبارات معيارية المرجع NRT:

ذكر بابام وهيوسك (Pohpam & Husek 1969) بأن هذه الاختبارات تعتمد عند تفسيرها للدرجات المستحصلة على خصائص المجموعة التي طبقت عليها حيث يجب أن تكون هذه الخصائص ممثلة مع خصائص الطلاب من حيث السن والنوع والمنطقة الجغرافية والمرحلة الدراسية، وتهتم هذه الاختبارات بالكشف عن الفروق الفردية بين الطلاب وذلك من خلال مقارنة أداء الطالب بأداء أقرانه كجماعة معيارية. بمعنى أن هذه الاختبارات تركز على وضع الطالب النسبي في علاقته بالآخرين.

الاختبارات محكية المرجع CRT

أشار بابام وهيوسك (Pohpam & Husek 1969) بأن هذه الاختبارات لا تعتمد عند تفسيرها لدرجة الطالب على خصائص مجموعته التي طبق عليها الاختبار، ولكنها تعتمد على مقارنة أداء الطالب بمحك أو مستويات أداء متوقعة ومحددة مسبقاً. ويتم تحديد هذه المستويات في ضوء الأهداف السلوكية (الإجرائية) المراد قياسها وتقيس هذه الاختبارات مدى تحقيق الطالب لهذه الأهداف السلوكية وبذلك يمكن التحقق من كفاءة كل طالب من اكتسابه للمهارات والمعارف المرجوة ومن ثم تشخيص نواحي الضعف والقوة.

أوجه الاختلاف بين الاختبارات معيارية المرجع ومحكية المرجع

أشار كثير من علماء القياس التربوي مثل جليزر، وإيبل، وبابام وهيوسك، وهامبلتون، (Hambleton et al, Papham & Husek, 1969; Ebel, 1979; Glaser, 1994) إلى وجود عدد من الاختلافات بين الاختبارات معيارية المرجع ومحكية المرجع وذلك في عدة نواحي منها مفهومها وبناءها وأهدافها وخصائصها والتي يمكن أن تقدم على النحو التالي:

1- الهدف الرئيس للاختبار:

يهدف الاختبار محكي المرجع إلى مقارنة أداء الطالب بمحك محدد مسبقاً وذلك لمعرفة الأهداف التي حققها والتي لم يحققها الطالب، أما الاختبار معياري المرجع فيهدف إلى مقارنة أداء الطالب بأداء مجموعته المعيارية من خلال تحديد عدد الأسئلة التي أجاب عليها الطالب بشكل صحيح.

2- استخدام الاختبار:

يستخدم الاختبار محكي المرجع للتأكد من تحقيق الطالب للأهداف السلوكية المحددة، أما الاختبار معياري المرجع فيعتمد على تزويدنا بمؤشرات عن درجة نجاح الطالب، وبالتالي إظهار الفروق الفردية بين الطلاب.

3- خصائص أسئلة الاختبار:

تتجمع الأسئلة في الاختبار محكي المرجع حول عدد محدد من الأهداف، بينما تنتشر الأسئلة في الاختبار معياري المرجع بشكل واسع حول نطاق الأهداف.

4- تفسير الأداء:

تفسر الدرجة في الاختبار محكي المرجع بناء على درجة القطع، بينما تفسر درجة الطالب في الاختبار معياري المرجع بناء على درجات معيارية تمثل موقع الطالب بين مجموعته الصفية والرتب المئينية.

5- بناء الفقرات:

يعتمد بناء الفقرات في الاختبار محكي المرجع على مدى تحقيق المفردة للهدف السلوكي التعليمي، بينما يعتمد بناءها في الاختبار معياري المرجع على تباين الدرجات ومعامل التمييز، ويتجنب اختيار المفردات الصعبة جداً والسهلة جداً.

6- التقويم:

يسعى الاختبار محكي المرجع إلى تحديد جوانب الضعف والقوة في كل من أداء الطلاب والبرامج التعليمية، بينما يسعى الاختبار معياري المرجع إلى تقويم أداء الطالب ومقارنته مع أداء الطلاب الآخرين.

تعريف الاختبارات محكية المرجع Definition of a CRT

أشار عدد من علماء القياس التربوي مثل بيرك، وجايغر، وهامبلتون ونوفيك Hambleton & Novick, 1973 Berk, 1982 Jaeger, 1989, إلى أن تعريف الاختبارات محكية المرجع يتعدد وفقاً لاختلاف آراء وخبرات علماء القياس التربوي. حيث توجد في أدبيات القياس التربوي أكثر من خمسين تعريفاً تصف الاختبارات محكية المرجع، وتتمحور تلك التعريفات حول النقاط التالية:

1- تعريف النطاق السلوكي content-domain الذي يقيسه الاختبار تعريفاً دقيقاً ذلك مما يساعد على اختيار مفردات تمثل هذا النطاق.

2- استخدام نتائج الاختبارات محكية المرجع لاتخاذ القرارات الخاصة بتحديد مستوى كفاءة الطالب مما يساعد على تشخيص نواحي ضعفه وقوته.

3- مقارنة أداء الطالب بمستويات أداء محددة مسبقاً (أهداف سلوكية محدد ة) يتخذ في ضوئها قرارات تربوية تتعلق بمدى إتقان الطالب أو عدم إتقانه للمهارات المراد قياسها.

طرق تحديد النطاقات السلوكية

تعتمد الطرق التقليدية في بناء الاختبارات المعيارية المرجع على وضع خطة تفصيلية تتضمن تحديدا للمحتوى المراد قياسه على البعد الأفقي لجدول مواصفات الاختبار، والأهداف المرتبطة بهذا المحتوى (غالبا ما تصاغ في عبارات عامة مثل معرفة المصطلحات، أو فهم المقروء، أو تطبيق المبادئ ...الخ على البعد الرأسي.

وتقدر أوزان تتناسب مع أهمية كل من المحتوى والهدف في كل خلية من خلايا الجدول. وتبنى مفردات الاختبار طبقا لهذه الأوزان.

أما الاختبار محكي المرجع يتطلب تحديدا أكثر دقة وتفصيلا للإطار السلوكي الذي يقيسه الاختبار قبل البدء الفعلي في كتابة المفردات. فالهدف من بناء هذا النوع من الاختبارات هو اتخاذ قرارات تعليمية تختلف عن القرارات التي تتخذ في حالة الاختبارات المعيارية المرجع.

فنحن هنا نهتم باتخاذ قرار عن مستوى أداء كل فرد بالنسبة إلى مجموعة من المهام أو المهارات أو المعارف التي تشكل النطاق السلوكي الذي يقيسه الاختبار. والحقيقة أن هذا النطاق يكون بمنزلة "المحك Criterion" الذي ينسب إليه أداء الفرد. فلكي نقول أن الفرد أجاب إجابة صحيحة عن 95٪ من المفردات التي يشتمل عليها الاختبار، يجب أن يكون المجتمع الشامل للمفردات محدد تحديدا واضحا. وتصبح المشكلة الأساسية في بناء الاختبارات محكية المرجع هي كيفية تحديد النطاقات السلوكية المراد قياسها والتي تنسب إليها درجة الفرد في الاختبار.
وفيما يلي عرضا لبعض الطرق في تحديد وصياغة النطاقات السلوكية المراد قياسها:

أولا: طريقة بابام في اعداد مواصفات الاختبار

يرى بابام Popham أن تحديد الأهداف التي يقيسها الاختبار وصياغتها بطريقة إجرائية (سلوكية) لا يعد كافيا لبناء الاختبار محكي المرجع. إذ يجب أن نعد مواصفات أكثر تفصيلا للنطاق السلوكي الذي يتضمن هذه الأهداف.

ولكن يمكن أن نسأل ما هي درجة الوضوح والدقة والنوعية اللازم توفرها في النطاق السلوكي حتى نعتبره معرفاً تعريفاً جيدا؟.

وفي الحقيقة لا توجد إجابات شافية لهذا السؤال. فالنطاقات السلوكية تختلف من موقف إلى آخر. فمثلا النطاق السلوكي في مجال الرياضيات ربما يشتمل على المهارات المختلفة اللازمة لحل مجموعة من المعادلات الآنية. والنطاق السلوكي في

مجال الاتجاهات ربما يشتمل على مجموعة من الاستجابات التي تتطلب الموافقة أو عدم الموافقة على عبارات تتصل بموضوع معين. وهذه المهارات تحتاج إلى تحليل دقيق لمكوناتها حتى نستطيع تحديد النطاق السلوكي الذي يشتمل عليها.

وفي أي من الأحوال يجب التأكد من أن صياغة النطاق السلوكي أصبحت من الدقة والوضوح بحيث يمكن استخدامها بسهولة في بناء الاختبار، ويجب أن يكون هناك اتفاق بين مجموعة من المحكمين على إمكانية كتابة مفردات متجانسة ومتسقة مع صياغة مكونات النطاق.

ونظرا لان الهدف من صياغة النطاق هو تحديد ما يشتمل عليه الاختبار أو ما يقيسه، فقد أطلق بابام على هذا المخطط اسم "مواصفات الاختبار Test Specification ولذلك تسمى هذه الطريقة بطريقة إعداد مواصفات الاختبار. وتشتمل هذه الطريقة على أربع مكونات تستخدم في تحديد ووصف السلوك الذي يمكننا من معرفة ما يستطيع الفرد أداؤه، ومالا يستطيع أداؤه في الاختبار. وينبغي أن نؤكد إن استخدام هذه الطريقة لا يمكننا من الحصول على نطاق سلوكي محدد تحديداً كاملا".

ويمكن التغلب على هذه المشكلة إلى حد ما بأن يحاول مصمم الاختبار معرفة حدود النطاق السلوكي المطلوب قياسه. وأن يقرر ما إذا كان من الأفضل اختيار قطاع عريض من السلوك ويحاول صياغته بالتفصيل لكي يحدد النطاق السلوكي تحديدا كافيا " أو اختيار قطاعات صغيرة من السلوك ويحدد مواصفات كل قطاع منها.

ويفضل بابام اختيار قطاعات صغيرة من السلوك لأنه يرى أنها أيسر في تحديدها وصياغتها. ولكنه يقترح مراعاة العوامل التالية عند اتخاذ قرار الاختيار بين نوعي القطاعات:

1- الفترة التعليمية: أي الفترة التي تستغرقها عملية التعليم لكي يكتسب الفرد المهارة أو السلوك المطلوب قياسه.

2- طرق القياس الممكنة: إذ أن السلوك يمكن قياسه بطرق متعددة.

3- تجانس فقرات الاختبار: فالنطاق السلوكي الذي يقيسه الاختبار يجب أن يسمح ببناء فقرات اختبار متجانسة، أي متشابهة في الشكل والمحتوى، وهذا لا يتأتى إلا إذا كان النطاق محدودا.

ولتحديد النطاق السلوكي الذي يقيسه الاختبار محكي المرجع أقترح بابام استخدام الخطوات أو المكونات الآتية :

أ- وصف عام لما يقيسه الاختبار:

وهو عبارة عن وصف عام للسلوك الذي يتم تحديده بالتفصيل. ويمكن اعتبار هذا الوصف هو صياغة للأهداف الإجرائية (السلوكية) التي يقيسها الاختبار إلا أن بابام يفضل تسميته "الوصف العام". وينبغي ملاحظة أن الوصف العام لما يقيسه الاختبار يمكن تعديله وتطويره بعد إجراء الخطوات التي تلي ذلك.

ب- عينة من فقرات الاختبار:

وهي توضح لمن يقوم ببناء الاختبار كيفية قياس الهدف قبل البدء في قراءة تفاصيل المواصفات التي سيأتي ذكرها في الخطوات التالية.

ج- تحديد عناصر المثيرات:

والمقصود بذلك تحديد ووصف المثيرات التي تشتمل عليها مفردة الاختبار. فمفردة الاختبار تعد بمثابة مثير يقدم للطالب لكي يستجيب له بطريقة معينة. لذلك يجب تحديد ووصف مكونات هذا المثير وتحليله إلى عناصره. والحقيقة أن "عناصر المثيرات" تعد من أهم مكونات مواصفات الاختبار، إذ يعتمد عليها الفرد الذي يقوم ببناء الاختبار اعتماداً أساسياً، وتحديد هذه العناصر يحتاج إلى كثير من الجهد والوقت. فهي تتطلب معرفة جميع العناصر الرئيسة التي يمكن الاستعانة بها في تكوين وصياغة مفردات الاختبار.

ويرى بابام وكرونلاند إنه إذا كان محتوى النطاق المراد قياسه غير محدود، يمكن

اختيار عينة ممثلة من هذا المحتوى وكتابة عناصرها الأساسية: فمثلا إذا كان المحتوى هو القصة الأدبية، ربما يكون من الصعب تحديد جميع أشكال وخصائص هذه القصص التي تشتمل عليها مثيرات مفردات الاختبار. لذلك من الأفضل تحديد عينة ممثلة لنطاق القصص الأدبية ولتكن عشرون قصة مثلا، أو أي عدد من القصص يمثل النطاق المطلوب. وهذا ييسر عملية تحديد العناصر التي تتضمنها المفردات.

د- تحديد عناصر الاستجابة:

والمقصود بذلك تحديد ووصف طريقة استجابة الفرد لعناصر المثيرات التي تشتمل عليها مفردة الاختبار. ويوجد نوعان رئيسيان من الاستجابات هما:

- الاستجابات المقيدة: مثل الاستجابات لمفردات الاختيار من متعدد أو الصواب والخطأ أو المزاوجة.
- الاستجابات المفتوحة: مثل الاستجابات لمفردات المقال والمفردات التي تتطلب إجابة قصيرة.

وتتكون عناصر الاستجابة من مجموعات من العبارات التي نحاول عن طريقها تحديد نوع وطريقة الاستجابة لمفردات الاختبار المستخدمة. فإذا كانت المفردة من نوع الاختيار من متعدد مثلا، فإنه يجب تحديد طبيعة الإجابة الصحيحة، وكذلك قواعد تكوين "البدائل" أي الإجابات غير الصحيحة بحيث يمكن أن يسترشد بها الفرد الذي يقوم بكتابة المفردات وبالمثل في الأنواع الأخرى من المفردات التي تتطلب إجابة مقيدة.

أما إذا كانت المفردة من النوع الذي يتطلب إجابة مفتوحة، فإنه يجب وضع محكات تفصيلية واضحة لطريقة الإجابة، وطريقة تصحيح الإجابات المختلفة. وهذا في الحقيقة ليس بالأمر اليسير، إذ إنه يتطلب إجراء عملية مسح للإجابات المتوقعة وتحديد عناصرها الرئيسة في عبارات واضحة.

ويشترط في كتابة الهدف السلوكي أن يرتبط بالهدف العام بحيث يحقق الشروط الآتية:

- يجب أن يمثل الهدف السلوكي عينة من السلوك التي يحتويها الهدف العام.
- يجب أن يتفق الهدف السلوكي والهدف العام في المستوى المعرفي.
- يجب أن يشتمل الهدف السلوكي على المكونات الرئيسة الأربع وهي: الدارس - الفعل السلوكي - شروط الأداء - مستويات الأداء.
- يجب أن لا تزيد عبارة الهدف الإجرائي عن (20) كلمة.

وتعد الإجابة صحيحة إذا حقق الهدف المحكات الأربعة السابقة وتعد الإجابة خاطئة إذا لم يحقق الهدف جميع المحكات أو حقق بعض منها فقط.

هـ- ملحق الموصفات:

يضاف أحياناً إلى المكونات السابقة ملحق للمواصفات تكتب فيه التفصيلات التي تؤدي إلى مزيد من التوضيح لمحتوى النطاق السلوكي أو محتوى مفردات الاختبار أو غير ذلك من عناصر المثير والاستجابة. وهذا يقلل من حجم التفاصيل التي تكتب في الخطوتين السابقتين وبذلك يمكن استخدام المواصفات بطريقة أكثر فاعلية.

ثانيا: صيغ المفردات Item Forms

وهذه الصيغة تتضمن قواعد تكوين المفردات أو هيكل أو تصميم يستخدم في تحديد المحتوى المراد قياسه، وبخاصة إذا كان المحتوى له طابع تكويني من مثل موضوعات الحساب والفيزياء وغيرها.

وقد توصل هايفلي Hively وأوزبرن Osburn إلى فكرة هذه الصيغ.

وفي هذه الصيغة يمكن أن نضمن تكوين مفردات متجانسة من حيث الشكل والمحتوى بحيث يمكن عن طريقها الحصول على النطاق الشامل للمفردات التي يتكون منها الاختبار.

ثالثا: مخططات تعتمد على القواعد اللغوية Linguistic-Based schemes

يرى بورموث Borsht 1970 إنه إذا اعتمد تفسير درجات الاختبارات محكية المرجع اعتماداً كليا على خصائص التعليم - وهو بالطبع الهدف الذي يبني من اجله هذا النوع من الاختبارات – فإن هذا يتطلب افتراض أن كل مفردة من مفردات الاختبار تكون مرتبطة ارتباطاً تاماً بالبرنامج التعليمي.

فالطرق التقليدية في بناء الاختبارات معيارية المرجع تسمح لكاتب الاختبار أن يغير من الصياغة والأسلوب اللغوي الذي استخدم في التعليم تبعا لما يراه. ومن الظواهر المؤكدة أن تغيير شكل المفردة أو صياغتها يكون له تأثير معين على مستوى صعوبة المفردة.

ونظراً لان كاتب المفردات يكون له حرية مطلقة في إجراء هذا التغيير، فإن صعوبة المفردات، وبالتالي الدرجات تتأثر بتحيز كاتب المفردات، وبالأسلوب اللغوي الذي أستخدمه في صياغتها. ولذلك فإن الاختبارات التي تبنى بالطرق التقليدية تقيس خصائص الشخص الذي يكتب المفردات إلى جانب قياسها لخصائص البرنامج التعليمي.

ولذلك أقترح بورموث أساسا لغويا لتحديد النطاق السلوكي الذي يستخدم في كتابة مفردات الاختبار أطلق عليها التعريفات الإجرائية Operational Definitions وهو يقصد بذلك "تحويلات المفردات Item Transformations".

وهو في هذا الصدد يذكر بورموث إنه يجب أن نتمكن من تطبيق هذه الإجراءات Operations بطريقة منتظمة في أي برنامج تعليمي إذا أردنا تكوين جميع المفردات التي تشتق من هذه الإجراءات. وتحقق هذا الشرط لا يضمن فقط إمكانية تعريف النطاق الشامل للمفردات، وإنما يضمن أيضاً تحييد الشخص الذي يقوم بتكوين هذه المفردات. أي أن لا يكون له أي تأثير على الصياغة اللغوية للمفردات. فمفردة الاختبار ترتبط بالتعليم إذا أمكن اشتقاقها بطريقة موضوعية

من جزء محدد من التعليم، وهذا لا يتأتى إلا إذا استخدمت مجموعة من الإجراءات التي يمكن تعميمها على البرامج التعليمية المختلفة.

وهذا يعني إنه يجب أن يكون هناك تناظر بين المفردة والجزء التعليمي المرتبط به، وكذلك بين الاستجابة الصحيحة للمفردة وذلك الجزء.

مثال:

في تعليم اللغة العربية، ربما تشير الإجراءات إلى "التحويلات" التي تجري على "الخصائص اللغوية للجملة" من مثل:

- "ركب الولد الحصان"، للحصول على المفردات:
- من ركب الحصان؟
- هل ركب الحصان الولد؟
- ماذا ركب الولد؟ ... وهكذا.

وبذلك يتضح من هذه الأمثلة إنه يمكن بناء القواعد التي تكون على أساسها مفردات الاختبار، بحيث نحصل على نفس المفردات أو الأسئلة مهما اختلف الأفراد الذين يقومون ببناء الاختبار. إذ أن هذه القواعد تتضمن الإجراءات التي تشتق على أساسها المفردات.

تصنيفات الاختبارات محكية المرجع Categories of CRT

تصنف الاختبارات محكية المرجع على حسب نطاقها السلوكي إلى قسمين وذلك ما أشار به نيتكو Nitko (1980) هما:

1. اختبارات محكية المرجع تعتمد على نطاقات مرتبة Ordered Domains

يعتمد هذا النوع على محكات تساعد على ترتيب النطاقات السلوكية وهذه المحكات عبارة عن:

أ- تحكيم المعايير الاجتماعية أو نوعية الأداء
(Judged Social or esthetic quality of performance)

ب- مستوى صعوبة أو تعقد المقرر الدراسي
Complexity or difficulty level of subject matter)

ج- درجة الكفاءة المطلوبة للمهارات المركبة
(Degree of proficiency with which complex skills are performed)

د- متطلبات اكتساب المهارات العقلية والحركية
Prerequisite sequence for acquiring intellectual and psychomotor skills

هـ- موقع السمات الكامنة المعرفة
((Location on empirically defined latiat trait

2. اختبارات محكية المرجع تعتمد على نطاقات غير مرتبة:

كما أشار نيتكو (Nitko 1980) إلى أن هذا النوع يتطلب تعريف النطاقات السلوكية، ويحتوي على أربعة أنواع من النطاقات يمكن إيجازها كالتالي:

أ- خصائص مثيرات النطاق وخطة اختيار مفردات الاختبار
(Stimulus Properties of Domain and Sampling Plan of Test)

ب- نطاقات تهتم بالخصائص اللفظية للمثيرات والاستجابات
(Verbal Statements of Stimuli and responses in domain)

ج- نطاقات تعتمد على تشخيص تصنيفات الأداء
(Diagnostic Categories of Performance)

د- نطاقات تعتمد على التجريد والسمات أو التكوينات الفرضية
(Abstractions, Raits or Constructs)

خطوات بناء الاختبارات محكية المرجع (الجانب المعرفي)

تستخدم الاختبارات محكية المرجع في تحديد مستوى أداء كل طالب بالنسبة إلى

نطاق سلوكي من المعارف والمهارات الأساسية المعرفة تعريفا دقيقا دون موازنة هذا الأداء بأداء غيره من الطلاب. فالنطاق السلوكي يكون بمثابة المحك Criterion الذي ينسب إليه أداء الطالب كما سبق أن ذكرنا. وهذا بالطبع يجعل تحديد النطاق السلوكي الذي يقيسه الاختبار تحديدا دقيقا هو المشكلة الأساسية التي تواجه من يقوم ببناء اختبار محكي المرجع.

ويمر بناء الاختبار محكي المرجع لقياس الجانب المعرفي بالخطوات الآتية:

أولا: تحديد المحتوى المراد قياسه

يعتمد ذلك على طبيعة وحدود ذلك المحتوى، فإذا كان المحتوى المراد قياسه محددا (مثل وحدة دراسية معينة، أو موضوع دراسي واضح المعالم) فإنه يمكن الاكتفاء بمعرفة مكونات هذه الوحدة أو هذا المحتوى. أما إذا كان المحتوى المراد دراسته واسعا وكبيرا فإنه من المستحسن تقسيمه إلى موضوعات فرعية مرتبطة ببعضها البعض بحيث يمكن قياسها كوحدة. ويتطلب ذلك بناء عدة اختبارات كل منها يرتبط بأحد هذه الموضوعات الفرعية.

ولا توجد قاعدة ثابتة لتقسيم المحتوى المراد قياسه، ولكن يجب أن يسمح هذا التقسيم ببناء اختبار أو اختبارات تكون مفرداتها بمثابة عينة ممثلة تمثيلا كافيا للموضوعات المختلفة التي يشتمل عليها المحتوى. كما يجب مراعاة طبيعة الطلاب المختبرين، والفترة الزمنية اللازمة لتعليم هذا المحتوى.

ثانيا: تحديد الأهداف العامة التي سيقيسها الاختبار

إن لكل محتوى أهداف مرتبطة به تهدف إلى قياس مدى تحققها لدى الطلاب. ومثال ذلك الهدف العام "يطبق الطالب الطريقة الاستنباطية في التفكير عند حل المسائل الرياضية".

وهذا الهدف العام يعبر عن نواتج متوقعة من عملية التعليم، ولكنها تحتاج إلى

أعادة صياغة حتى تعبر عن أنشطة سلوكية تكون دليلا على أن الطالب قد طبق الطريقة الاستنباطية وما إلى ذلك.

ثالثا: تحليل الأهداف العامة إلى أهداف جزئية

ونعني بذلك إعادة صياغة الأهداف العامة عن طريق وصف عينة من الأهداف السلوكية التي يمكن أن تتخذ دليلاً على تحقق كل هدف من الأهداف العامة. وعادة يواجه مصمم الاختبار بعض الصعوبات عند إجراء هذا التحليل. ويعتمد هذا على النطاق السلوكي المراد قياسه. فإذا كان هذا النطاق محدودا فإنه يمكن استخدام طريقة جاجني Gagne 1968 في تحليل المهام Task Analysis، أو طريقة ميجار Mager 1972 في تحليل الأهداف Goal Analysis. وهاتان الطريقتان تفيدان في التوصل إلى مجموعة الأهداف أو المهام التي تسهم في تحقيق الهدف العام.

رابعا: تحديد مواصفات النطاق السلوكي الذي يقيسه الاختبار

عند الحصول على الأهداف الإجرائية من التحليل السابق قد لا تكون واضحة بدرجة تسمح بتحديد نطاق مفردات الاختبار التي تقيس السلوك المطلوب. وبالطبع يصعب اختيار عينة ممثلة لمفردات الاختبار من النطاق إذا كان هذا النطاق الذي يقيس هدفا معينا غير واضح. وقد سبق وأن ذكرنا أن أداء الفرد في عينة من مفردات الاختبار التي تقيس هدفا معينا يستخدم في تقدير درجة إتقانه للنطاق الشامل للمفردات التي تقيس هذا الهدف. لذلك فإنه يجب تعريف وتحديد هذا النطاق الشامل للمفردات تحديدا جيدا يمكن سحب عينات ممثلة من هذه المفردات.

فإذا كان النطاق الشامل يصعب تحديده تحديدا كاملا، فإنه يمكن إعداد مواصفات الاختبار باستخدام طريقة بابام التي تم عرضها في تحديد النطاقات السلوكية. أما إذا كان النطاق الشامل حدوده معرفة تعريفا جيدا فإنه يمكن استخدام صيغ المفردات التي اقترحها هايفلي التي اشرنا لها سابقاً.

خامسا: تكوين مفردات الاختبار

وتشمل الخطوات الآتية:

1- اختيار أنواع المفردات المناسبة لقياس الأهداف

بعد تحديد مواصفات النطاق السلوكي يجب اختيار أفضل أنواع المفردات التي تقيس السلوك الذي تم تحديده في المواصفات الخاصة بكل هدف. فمن المعلوم أن بعض أنواع المفردات تصلح لقياس أهداف معينة بدرجة أفضل من غيرها. فمثلا تصلح مفردات الاختيار من متعدد، ومفردات المزاوجة، والصواب والخطأ، لقياس التذكر والفهم وأحيانا التطبيق، بينما تصلح مفردات المقال لقياس القدرة على تنظيم المعلومات، والاستنتاج، والتفسير، وإعادة صياغة الأفكار وهكذا.

2- تحديد العدد المناسب من المفردات

في هذه الخطوة نحاول تحديد العدد المناسب من المفردات التي تقيس الأهداف الممثلة في النطاق السلوكي الذي سبق تحديده ويجب مراعاة العوامل الآتية:

ذكر كل من هامبلتون وإيقنور Hambleton & Eignor (1980) بأن هناك عدة عوامل ينبغي وضعهما في الاعتبار عند تحديد أسئلة الاختبار وهما:

1- **العلاقة بين عدد أسئلة الاختبار والأهمية الخاصة بكفاية معينة:** تتفاوت أهمية الكفايات المراد قياسها في الاختبار، لذا ينبغي أن يختار مصمم الاختبار عند قياسه كفايات متعددة أسئلة تتناسب مع أهمية النطاق المراد قياسه.

2- **العلاقة بين عدد أسئلة الاختبار ومستوى الحد الأدنى المقبول لدرجة ثبات الاختبار:** ففي حالة الاختبارات معيارية المرجع تستخدم المعادلة العامة لسبيرمان براون لتقدير معامل الثبات إذا زاد أو نقص طول الاختبار (ن) من المرات. أما في حالة الاختبارات محكية المرجع فقد اقترح هامبلتون Hambleton بتحديد قيمة الدقة المطلوبة ((Degree of Precision) في تقدير درجة الطالب في النطاق، ويتم ذلك وفق الصيغة التالية:

$$\text{طول الاختبار} = \frac{0,25}{(\text{درجة الدقة المطلوبة})^2}$$

حيث (0,25) مقدار ثابت.

وأشار كل من هامبلتون، وآخرون (Hambleton et al 1978) بأن الهدف من تطبيق الاختبار محكي المرجع هو تصنيف الطلاب إلى فئتين متقنة وغير متقنة، لذا فإن تحديد طول الاختبار المناسب الذي يقلل من أخطاء التصنيف (Number of Classification Errors) يتطلب تطبيق اختبار طويل يقلل من أخطاء القياس.)

وأضاف كل من هالادين ورويد Haladyn & Roid بأن قيمة أخطاء التصنيف تقل بشكل عام عند زيادة طول الاختبار، لكن النقص في قيمة الخطأ يكون ملحوظاً عند الانتقال من اختبار طوله (10) فقرات إلى اختبار طوله (20) فقرة، وأما النقص في قيمة الخطأ عند الانتقال من اختبار طوله (20) فقرة إلى اختبار طوله 30-40 فقرة يكون صغيراً نسبياً لدرجة قد لا تبرر استخدام فقرات تزيد عن (20) فقرة.

3- **العلاقة بين عدد المفردات وزمن تطبيق الاختبار:** يعتمد عدد المفردات بلا شك على الزمن اللازم لتطبيق الاختبار، ويعتمد هذا بالتالي على نوع مفردات الاختبار المستخدمة، وعلى خصائص الأفراد المختبرين.

3- كتابة مفردات الاختبار

وهذه الخطوة تحتاج إلى عناية كبيرة. فمفردات الاختبار محكي المرجع تبنى على أساس مواصفات النطاق السلوكي الذي سبق أعداده. لذلك يجب أن تقيس هذا النطاق بدرجة كبيرة من الدقة، وان يكون مستوى صعوبة كل مفردة مناسبا لمستوى صعوبة الهدف الذي تقيسه، ولمستواه المعرفي. وان تكون عينة المفردات ممثلة للنطاق السلوكي للأهداف. كما يجب أن تراعى الأصول الفنية في كتابة الأنواع المختلفة للمفردات.

بناء الاختبارات محكية المرجع لقياس الجانب الوجداني

يحتاج بناء المقاييس والاختبارات محكية المرجع في الجانب الوجداني إلى قدر كبير من المهارة. فتقويم الجانب الوجداني للسلوك يعد أكثر صعوبة من تقويم الجانب المعرفي لأسباب كثيرة نذكر منها:

1- صعوبة تحديد وتعريف مكونات الجانب الوجداني للسلوك. وكذلك صعوبة ملاحظة هذه المكونات غير المنظورة يحتاج إلى تحليل سلوكي، وملاحظة سلوك الفرد ملاحظة مباشرة تعكس هذه المكونات. وبالطبع يواجه مصمم الاختبار نفس المشكلة عند قياس مكونات الجانب المعرفي، أو ما يسمى بالسلوك أو الأداء الأقصى حيث يؤدي الفرد أفضل أداء ممكن ليحصل على درجة مرتفعة في الاختبار. ولكن المشكلة تزداد حدتها عند قياس الجانب الوجداني للسلوك أو ما يسمى بالسلوك أو الأداء المميز، حيث يستجيب الفرد للمواقف الاختيارية كما هو بالفعل في سلوكه اليومي المعتاد.

2- يتطلب تقويم الجانب الوجداني للسلوك ملاحظة الفرد ملاحظة متكررة في عينات مختلفة من المواقف التي يمكن أن تظهر فيها الاستجابات التي نهتم بملاحظتها وتقويمها. فلكي ندرس ميول فرد، فأننا نحتاج إلى ملاحظته في وقت فراغه مثلاً، ولكن ملاحظة سلوكه في موقف واحد لا يكفي للحكم على أدائه المتميز. فالسلوك الوجداني يعتمد على مزاج الفرد، والألفة بالموقف، وتفاصيل البيئة المحيطة به، وعوامل أخرى كثيرة. فالفرد الذي يبدو عدوانياً في كثير من المواقف والظروف ربما نحكم عليه بأنه شخص عدواني، فقد يكون في حالة من الضيق والتوتر نتيجة قلقه المستمر على شيء ما، ولكنه ربما يبدو في حالة تكيف تام بعد انقضاء عدة أسابيع أو عدة شهور. ولذلك فإن ما نلاحظه من سلوك فرد معين في مواقف قليلة، ربما يكون مجرد انحراف وقتي، ولكن بلا شك أن هذا السلوك قد ميز الفرد في هذا الوقت.

ونظرا لصعوبة ملاحظة الفرد طوال اليوم أو خلال أطول فترة ممكنة، فأننا

نعتمد على الاستجابات اللفظية التي نطلبها منه، عن طريق الأسئلة المباشرة أو عن طريق الاستبيانات. ولكن يجب تفسير البيانات المستمدة من هذه الاستبيانات بحذر شديد، والتأكد من صدق التفسير باستخدام ما نجمعه عن الفرد من ملاحظات أثناء قيامه بنشاطات مختلفة.

3- نظرا لصعوبة التحليل السلوكي الدقيق لكثير من المفاهيم الوجدانية من مثل مفهوم الذات أو مفهوم القيم، أو مفهوم الاتجاهات، فأننا لا نستطيع الاعتماد على أداة قياس واحدة تقيس أبعاد كل مفهوم من هذه المفاهيم، إذ من الصعب احتواء هذه الأبعاد جميعا وقياسها مباشرة بأداة قياس واحدة. لذلك ربما يكون من الأفضل اتباع إستراتيجية الالتفاف حول هذه الأبعاد، أي نستخدم عددا من المقاييس المختلفة التي تحتوي أهم هذه الأبعاد، وبذلك نحصل على صورة أفضل عن مكونات كل من هذه المفاهيم.

4- يحاول الأفراد في كثير من الأحيان تزوير استجاباتهم أو سلوكهم المطلوب قياسه إذا ما شعروا أنهم سوف يكونون عرضة للنقد، أو أن الدرجات التي ستقدر لهم في المواقف الاختبارية سوف تؤثر بشكل من الأشكال على التحاقهم بعمل معين أو على مستقبلهم الدراسي مثلا. ولذلك يجب أن نفرق بين استخدام البيانات المستمدة من استجابات الأفراد لهذه المواقف لغرض التقويم الجماعي، واستخدامها لغرض التقويم الفردي. ففي الحالة الأولى يمكن أن يستجيب الفرد لهذه المواقف (التي تكون عادة استجابات لفظية) دون الحاجة إلى ذكر أسمائهم على صحيفة الاستبانة مثلاً. ومن هنا يمكن الاستفادة من البيانات في اتخاذ قرارات جماعية مثل تأثير موضوع أو قضية معينة في السلوك الوجداني لمعظم الأفراد. أما في الحالة الثانية فإنه يمكننا أن ننصح الأفراد بكتابة أسمائهم على صحيفة الاستبانة أو غيرها من التقارير الذاتية حتى يتمكن كل منهم من تقويم نفسه تقويماً ذاتيا " بالنسبة للبعد الوجداني المطلوب قياسه.

ويفضل في مثل هذه الحالة استخدام الاستبيانات قهرية الاختيار للتغلب على

عامل الرغبة الاجتماعية أو الميل إلى المعايير الاجتماعية، وهو نوع من تحيز الاستجابات، حيث يحاول الفرد بطريقة شبه شعورية أن يكون مقبولا اجتماعياً. لذلك فهو يوافق على عبارة مثل "أحاول قدر استطاعتي تجنب إثارة المشاكل" أو "أنا شخص ذكي".

وقد أكد الين ادوردز E. Edwards وهو من علماء النفس الذين اهتموا ببحث هذا النوع من تحيز الاستجابات، أن موافقة الفرد على مثل هذه العبارات لا يعتبر نوعا من الكذب لأن الاستجابة تكون شبه شعورية في حين أن الكذب يكون متعمدا. ويعتمد الفرد في استجابته على محتوى العبارات، ويكون مدركاً للعلاقة بين محتوى العبارات وطريقة تصحيحها. وللتغلب إلى حد ما على هذه المشكلة اقترح ادوردز استخدام هذا النوع من الاستبيانات، وفيه تعرض بنود الاستبانة على الفرد بحيث يشتمل كل بند منها على عبارتين تبعد كل منهما عن المعايير الاجتماعية بنفس الدرجة، وعلى الفرد أن يحدد قبوله أو رفضه لإحداهما.

فمثلاً يكون البند مشتملا " على عبارتين كالتالي:

أ- أحب أن أتحدث عن نفسي أمام الناس.

ب- أحب أن اعمل لتحقيق الهدف الذي اصبوا إليه.

فالفرد الذي يختار الاستجابة الأولى مثلاً ربما يقدر له درجة واحدة على ميزان أو مقياس الاستعراضية Exhibition Scale، أما إذا اختار الاستجابة الثانية، فإنه ربما يقدر له درجة واحدة على ميزان أو مقياس التحصيل والإنجازAchievement Scale.

ويمكن باستخدام هذا النوع من الاستبيانات للتغلب بدرجة معقولة على مشكلة الميل إلى المعايير الاجتماعية في استجابات الأفراد للاستبيانات والقوائم Inventories.

وقد أقترح كراندال Crandall وكاتكوفسكي Katkovsky 1965 طريقة أخرى تفيد في التعرف على العبارات المرغوبة اجتماعياً. وتتمثل هذه الطريقة في انتقاء مجموعة من العبارات التي تؤدي إلى استجابات مرغوبة اجتماعياً وأضافتها إلى

عبارات المقياس المطلوب. وتعتبر عندئذ هذه العبارات المنتقاة مقياساً للميل إلى المعايير الاجتماعية.

فاستخدام هذا المقياس الأخير يمكننا من التعرف على عبارات المقياس الوجداني التي ترتبط ارتباطاً عالياً بدرجة الفرد في الميل إلى المعايير الاجتماعية. وبذلك يمكن حذف هذه العبارات من المقياس المطلوب.

تحليل الجوانب الوجدانية للسلوك:

من هذا يتضح أن قياس الأبعاد المختلفة للسلوك الوجداني باستخدام الاختبارات والمقاييس محكية المرجع، يتطلب التفكير في العوامل التي تؤثر في الطريقة المميزة لاستجابات الفرد للمواقف المختلفة المرتبطة بهذه الأبعاد، أي في النزعة العامة لاستجاباته.

وبالطبع لا نستطيع قياس هذه النزعة بدرجة تامة من التأكد، وإنما يمكن تقديرها. فمن المستحيل التنبؤ بدرجة تامة من الثقة بكيفية سلوك الفرد في ظروف أو مواقف معينة، وإنما يمكن معرفة احتمال أن يستجيب بطريقة معينة في بعض المواقف.

ويكمن الفرق الرئيسي بين قياس كل من المجالين المعرفي والنفسحركي من جهة، والمجال الوجداني من جهة أخرى، في الفرق بين ما يستطيع الفرد أن يفعله، وبين ما سيفعله في المستقبل. بمعنى إنه في الجانبين المعرفي والنفسحركي نفترض أن الفرد ربما لا يكون قادرا بالفعل على تحقيق هدف معين، بينما في الجانب الوجداني نفترض أن الفرد يمكنه تحقيق هدف وجداني، ولكنه ربما لا يفضل أن يسلك السلوك المرتبط بالهدف بمحض إرادته.

ولذلك فإن قياس الجوانب الوجدانية للسلوك يتطلب الحصول على دلائل عن الأفراد أما في صورة نواتج معينة، أو في صورة سلوك يمكن ملاحظته ملاحظة مباشرة في بعض المواقف. وهذه المواقف التي تعتبر بمنزلة المثيرات التي نلاحظ عن

طريقها النتاج أو السلوك المعين يجب أن تسمح لنا بالاستدلال على مشاعر الأفراد في مواقف أو مثيرات مستقبلية مشابهة لها. فمثلاً ربما يحاول معلم قياس توافق طالب معين مع أقرانه عن طريق ملاحظة سلوكه في مواقف تتطلب علاقات مع الأقران، أو عن طريق استخدام التقارير الذاتية على أمل أن يحصل المعلم بهذه الطرق على بعض الدلائل التي تساعده في التنبؤ بسلوك الطالب مع الآخرين في مستقبل حياته.

فعلى الرغم من اختلاف المواقف التي يمكن أن يستخدمها المعلم في هذا التقويم عن مواقف الحياة المستقبلية للطالب، إلا إنه يجب أن يكون هناك نوع من التطابق أو التشابه بين هذه المواقف. فهذا التشابه هو الذي يسمح بالتنبؤ بقدر من الثقة بسلوك الطالب في المستقبل. وبالطبع يرتبط هذا ارتباطاً وثيقاً بطرق بناء أدوات قياس الجوانب الوجدانية. إذ أن هذه الأدوات تشتمل عادة على مواقف سلوكية ترتبط بالأبعاد الوجدانية المطلوب قياسها. وهذا يتطلب بالطبع تحليل هذه الأبعاد تحليلاً سلوكياً أي تحديد النطاق السلوكي تحديداً جيداً. ويجب أن يعتني بتحديد هذا النطاق السلوكي بالقدر نفسه من الدقة الواجبة في الاختبارات محكية المرجع التي تقيس الجوانب المعرفية.

طرق تقدير صدق الاختبارات محكية المرجع:

1- الصدق الوصفي Descriptive - Validity

ذكر بابام Popham 1978 بأن الصدق الوصفي عبارة عن محاولة لتأكيد مدى ما يقيسه فعلاً الاختبار محكي المرجع، ويعتبر تحديد الصدق الوصفي شرط أساسي لأنواع الصدق الأخرى وذلك لأنه أفضل وسيلة لتفسير أداء الطالب في الاختبار. ويمكن تحديد الصدق الوصفي باستخدام أساليب مختلفة، ومن ذلك الاستعانة بمجموعة من المتخصصين لتحديد مدى مطابقة Congruent مفردات الاختبار مع المجال السلوكي المراد قياسه.

2- الصدق الوظيفي Functional - Validity

أشار بابام (Popham 1978) بأن المقصود بالصدق الوظيفي هو دقة الاختبار محكي المرجع في تحقيق الغرض) الوظيفة (الذي صمم من أجله الاختبار، أي إنه لا يقتصر على وصف أداء الطالب في الاختبار، ويستخدم الصدق الوظيفي عندما يكون الهدف من تطبيق الاختبار وصف وتحديد الوظيفة المراد تحقيقها من الاختبار، لذا فإن الصدق الوصفي يعتبر شرطاً لازماً لحساب الصدق الوظيفي. ويقابل الصدق الوظيفي في القياس المعياري المرجع صدق المحك الخارجي Related criterion Validity والذي يدل على قدرة الاختبار على التنبؤ بأداء طالب ما في موقف تالي من أداء الطالب على اختبار سبق التأكد من صدقه، وأطلق على هذا النوع من الصدق (بالصدق الوظيفي) بدلاً من صدق المحك الخارجي لأن بعض الوظائف المراد قياسها في الاختبار محكي المرجع تتطلب محكاً خارجي والبعض الآخر لا يتطلب ذلك.

وأضاف هامبلتون بأنه يمكن تحديد صدق القرار بمقارنة أداء مجموعتين محكيتين أو أكثر بدرجة قطع محددة. كما يمكن تحديد صدق القرارات التعليمية المبنية على درجات الاختبارات محكية المرجع بمقارنة أداء مجموعتين (الأولى تلقت التعليم والثانية لم تتلق التعليم) بدرجة قطع الاختبار. ويمكن تقييم صدق القرار لكل هدف تعليمي بجمع النسب المئوية للطلاب المتقنين الذين اجتازوا درجة القطع والنسب المئوية للطلاب غير المتقنين الذين لم يجتازوا درجة القطع، وذلك بهدف تحديد الطلاب المتقنين فعلاً والطلاب غير المتقنين فعلاً. ويمكن تحديد صدق القرار أيضاً بحساب معامل الارتباط بين متغيرين ثنائيين الأول متغير الانتماء إلى المجموعة (Group Membership) والثاني متغير قرار الإتقان. (Mastery Decision)

3- صدق انتقاء النطاق السلوكي Domain Selection Validity

أشار بابام (1978) Popham بأن صدق انتقاء النطاق يتعلق بمدى دقة باني

الاختبار في اختيار نطاق سلوكي محدد يمكن أن يستخدم كمؤشر ((Indicator) لإيضاح مستوى الطالب بالنسبة إلى (نطاق عام) يعد أكثر عمومية. ويمكن تحديد صدق انتقاء النطاق السلوكي بأساليب متعددة منها ما يلي:

أ- تدريس مجموعات صغيرة من الأفراد المهارات التي يشتمل عليها كل نطاق من النطاقات السلوكية حتى يصلوا إلى درجة الإتقان، يلي ذلك قياس إمكانية تعميم إتقان الطلاب لنطاق واحد تجاه النطاقات الأخرى.

ب- مقارنة أداء مجموعتين من الأفراد إحداهما (تلقت التعليم) والأخرى (لم تتلق التعليم)، وذلك بتطبيق عدد من الاختبارات، ثم يحدد النطاق السلوكي الذي يميز (Discriminates) بصورة أفضل بين المجموعتين.

طرق تقدير ثبات الاختبارات محكية المرجع:

تختلف وتتعدد طرق تقدير ثبات الاختبارات محكية المرجع باختلاف استخداماتها، وصنف هامبلتون، وآخرون (Hambleton, et al 1978) تلك الطرق إلى ثلاث مجموعات رئيسة، وستطرق لمجموعتين هما: تقدير ثبات درجات الطلاب في النطاق السلوكي وتقدير ثبات تصنيف الطلاب في مجموعات على حسب درجة تمكنهم من النطاق السلوكي.

أولا- تقدير ثبات درجات الطلاب في النطاق السلوكي:

طرق هذه المجموعة تهتم بتقليل خطأ التباين الناتج عن الفروق بين درجات الطلاب في النطاق السلوكي بالإضافة إلى درجات النطاق، ومن تلك الطرق ما يلي:

معامل ليفنجستون Livingston-Index

ذكر كل من كروكر، والجاينا Algina & Crocker بأن هذا المعامل يهتم بإيجاد انحراف درجات الأفراد عن درجة القطع، ويهتم هذا المعامل أيضاً بمفهوم متوسط مجموع مربعات انحرافات الدرجات الملاحظة المتوقعة وذ لك بحساب قيمة انحراف درجات كل طالب عن درجة القطع في الاختبار، ويتم ذلك وفق الصيغة الرياضية التالية:

$$K^2 (X,T)= \frac{6^2T+(Mt-N \sim C)^2}{6^2X+(Mx-N \sim C)^2}$$

حيث أن:

$K^2 (X,T)$: ترمز إلى معامل ليفنجستون.

6^2T: ترمز إلى متوسط مجموع مربعات انحرافات الدرجات الملاحظة حول درجة القطع (C).

6^2X: ترمز إلى متوسط مجموع مربعات انحرافات الدرجات الحقيقية عن درجة القطع (C).

M: ترمز إلى متوسط درجات الطلاب في النطاق السلوكي Domin الذي يقيسه الاختبار.

$N \sim$: ترمز إلى عدد الأسئلة

C : ترمز إلى درجة القطع

ولتقدير معامل ليفنجستون $K^2 (X,T)$ في حالة استخدام اختبارين لكل طالب تستخدم الصيغة الرياضية التالية:

$$K\hat{}^2 (X,T)= \frac{P\hat{}xx`\ 6\hat{}x\ 6\hat{}x` + (M\hat{}x - N \sim C)(M\hat{}x` - N \sim C)}{\sqrt{\{6\hat{}x^2 + (M\hat{}x - N \sim C)^2\}\{6\hat{}x`^2 + (M\hat{}x` - N \sim C)^2\}}}$$

حيث أن:

$P^{\hat{}}xx`$: ترمز إلى معامل الارتباط بين درجات كلا الاختبارين.

$6^{\hat{}}x^2$: ترمز إلى متوسط مربعات انحرافات الدرجة الحقيقية عن درجة القطع.

$M^{\hat{}}$: ترمز إلى متوسط الدرجات الخام.

N : ترمز إلى عدد الأسئلة.

C : ترمز إلى درجة القطع.

ثانيا- تقدير ثبات تصنيف الطلاب في مجموعات على حسب درجة تمكنهم من النطاق السلوكي:

تهتم طرق هذه المجموعة بتحديد الأخطاء الناتجة عن عدم اتساق التصنيف عند تطبيق اختبارين متوازيين أو في حالة إعادة تطبيق الاختبار على أفراد العينة، ومن تلك الطرق ما يلي:

1- طرق تتطلب تطبيق الاختبار مرة واحدة:

أ- طريقة هاريس Harris Method

أشار هاريس (Harris 1974) بأن هذه الطريقة لا تعتمد في تحديدها على متغير طول الاختبار (k) وإنما تقوم على الارتباط بين متغير يمثل مجموع درجة الطالب في الاختبار ومتغير ثنائي آخر يمثل تصنيف الطالب إلى متقن أو غير متقن على ضوء درجة القطع، لذا فإنه يتم تصنيف الطلاب إلى أربع مجموعات في (جدول رباعي) وهي طلاب متقنين فعلاً، و طلاب غير متقنين فعلاً. ومجموعة طلاب الخطأ الأول (α) ومجموعة طلاب الخطأ الثاني (β) ويتم تحديدهم بناء على درجة قطع الاختبار. ويطلق على هذه الطريقة بمعامل (هاريس) ويرمز له بالرمز (Mc) ويتم تقديره بالصيغة التالية:

$$M_c = \frac{SSb}{SS_b + SS_w}$$

حيث إن:

M_c = معامل ثبات هاريس.

SS_w ، SS_b = عبارة عن مجموع المربعات (داخل، بين) المجموعات.

وتمتد قيمة المعامل ما بين (صفر وواحد).

ب- معامل كابا (هاينا) Huynh Kappa Coefficient

يمكن تقدير معامل (هاينا) في حال تطبيق اختبار واحد أو اختبارين متوازيين، ونظراً لصعوبة تقدير هذا المعامل في حال تطبيق اختبارين متوازيين، لذا سيتم التطرق لكيفية تقدير معامل الثبات (هاينا) لاختبار واحد فقط، وفي هذا الصدد أشار سابكوفياك Subkoviak بأن تقدير هذا المعامل يستند إلى عدة فروض هي:

أن يتمثل توزيع درجات الطلاب الحقيقية في الاختبار (Beta. Distribution) شكل بيتا - إذا طبق عدد من الاختبارات على طالب ما، فإنه يفترض أن يتخذ توزيع درجاته في الاختبار شكل التوزيع ذي الحدين، وفي هذه الحالة يجب أن يكون تقدير درجات أسئلة الاختبار أما صفر أو واحد، وأن تكون الإجابة على الأسئلة مستقلة إحصائياً بحيث لا تؤثر الإجابة على إحدى المفردات على إجابة المفردة الأخرى، وأن تتساوى تقريبا معاملات صعوبة جميع أسئلة الاختبار.

ويتم تقدير معامل ثبات(هاينا) وفق إحدى الصيغتين التاليتين:

$$K = \frac{P\hat{}zz - P^{2}\hat{}z}{P\hat{}z - P^{2}\hat{}z}$$

$$P\hat{}o = 1 - 2(P\hat{}z - P\hat{}zz)$$

ويتطلب تحقيق الفروض السابقة إلى إجراء الخطوات التالية:

- تقدير متوسط درجات الطلاب ^M والانحراف المعياري 6^2 ومعامل ثبات كودر ريتشاردسون 21 (a^21) وتحديد درجة القطع (C) وبارمتر توزيع الفا (α).

- تحويل القيم السابقة إلى بيانات معيارية وفق الصيغ الآتية normalize data

$$M\hat{} = \sin^{-1}\sqrt{M\hat{}/n}$$

M^ = ترمز إلى المتوسط

n = ترمز إلى عدد المفردات

$$6 = \sqrt{(a\hat{}21+1) + (a\hat{} + N)}$$

6`^ = ترمز إلى الانحراف المعياري

$$P\hat{} = a\hat{}21\sqrt{(n-1) / (n + a\hat{}21)}$$

P^ = ترمز إلى معامل الثبات

$$C\hat{} = \sin^{-1}\sqrt{(C - 0.5)\, n}$$

C^ = ترمز إلى درجة القطع

تقدير الانحراف المعياري ((normal deviate بناءاً على درجة القطع (Z)

$$Z = (C` - M) / 6`\hat{}$$

تقدير قيمة الاحتمال (P^Z) من الجداول الخاصة بذلك بحيث تكون هذه القيمة اقل من قيمة (Z).

ثم تحديد قيمة الاحتمال (P^ZZ) من جداول دالة التوزيع الثنائي الاعتدالي بحيث تكون قيمة الاحتمال أيضاً اقل من قيمة (Z).

2- طرق تتطلب تطبيق الاختبار مرتين:

أ- معامل كارفر Carver Method

أشارت سابكوفياك Subkoviak إلى أن معامل كارفر يعد من الطرق الأولى المقترحة

لتقدير ثبات الاختبارات محكية المرجع التي تتعلق باتساق قرار التصنيف. وتقوم هذه الطريقة على تطبيق اختبارين متوازيين على مجموعة واحدة من الأفراد، يليه مقارنة نسبة عدد الطلاب الذين تم تصنيفهم ضمن المتقنين في كلا الاختبارين، فإذا كانت النسبتين متماثلتين أو قريبتين من بعضهما فبذلك يعتبر الاختبارين ثابتين، ويعاب على هذه الطريقة بأنها غير حساسة لاتساق قرارات كل طالب في الاختبارين.

ويمكن إيجاد قيمة معامل (كارفر) وفق الجدول التالي:

		الاختبار (ب)	
		غير متقن	متقن
الاختبار (أ)	متقن	ب	أ
	غير متقن	ج	د

ويحسب معامل كارفر باستخدام الصيغة التالية:

$$= \frac{أ + ج}{ن}$$

حيث إن:

أ = مجموع الطلاب المتقنين بالفعل.

ج = مجموع الطلاب غير المتقنين بالفعل

ن = أ+ ب+ ج+ د

ب- معامل (كابا) Swaminathan- Hambleton - Algina

ذكر كل من هامبلتون، وآخرون Hambleton, et al 7819 بأن هذه الطريقة تعتبر امتداداً لطريقة (كارفر) السابقة حيث تعتمد على فرق النسبة بين الطلاب المتمكنين وغير المتمكنين في صورتي الاختبار المتوازيتين أو في مرتي تطبيق الاختبار، وتمت د قيمة معامل الثبات بين (+1، -1).

ويمكن تقدير معامل الثبات وفق الصيغة التالية

$$Po\hat{} = Sum\ (P\hat{}\ KK)$$

حيث إن:

$Po\hat{}$ ترمز إلى معامل الثبات أي اتساق التصنيف.

$P\hat{}\ KK$ ترمز إلى نسبة الطلاب الذين تم تصنيفهم ضمن المتقنين في المستوى K في مرتي تطبيق الاختبار.

M ترمز إلى عدد مستويات الإتقان.

ونظراً لأن هذه الطريقة لا تأخذ بعين الاعتبار اتساق التصنيف الناتج عن عامل الصدفة، لذا أقترح كل من سواميناثان وهامبلتون وألجاينا استخدام معامل كابا(Kappa Coefficient) الذي ينسب إلى كوهن.

$$K = \frac{Po\hat{} - Pc\hat{}}{1 - pc\hat{}}$$

حيث أن:

K = معامل كابا

$Po\hat{}$ = نسبة الاتفاق الملاحظ في التصنيفات وتحسب بالصيغة السابقة

$$Po\hat{} = sum\ (PKK)$$

$pc\hat{}$ = ترمز إلى نسبة الاتفاق المتوقع في التصنيفات.

وتحسب بالصيغة التالية:

$$Po\hat{} = sum(PK.\ P.\ K)$$

حيث إن:

(PK.) (P. K) = نسب الطلاب المصنفين ضمن فئة المتقنين في المجموعة (K) في مرتي تطبيق الاختبار على التوالي.

مسميات وتعريف درجة القطع Definitions of cut off score

مسميات درجة القطع:

تتعدد مسميات درجة قطع الاختبارات بتعدد آراء وخبرات علماء القياس والتقويم التربوي في مجال الاختبارات محكية المرجع، ومن تلك المسميات ما ذكرها هالبن، وآخرون Halpin, et al 1983 وهي كما يلي:

- درجة القطع Cut off Score
- الحد الأدنى من الكفاية Minimum Competency
- درجة المحك Criterion Score
- درجة الاجتياز Passing Score
- مستوى الحد الأدنى من الكفاية Level of Minimal Competency

تعريف درجة القطع:

نظراً لأهمية مفهوم درجة القطع في القياس محكي المرجع، الأمر الذي أدى إلى تعدد تعريفات علماء القياس التربوي والنفسي لمفهوم درجة القطع، ومن تلك التعريفات ما أشار إليه هامبلتون (Hambleton, 1978) حيث عرف درجة القطع "بأنها نقطة على متصل درجات الاختبار تستخدم لتصنيف الطلاب إلى فئتين تعكس مستويات الأداء المختلفة بالنسبة لهدف معين أو (الأهداف) المراد قياسها في الاختبار".

كما عّرف بيرك Berk درجة القطع بأنه "تلك الدرجة التي تزيد من نسبة التصنيفات الصحيحة أو في المقابل تقلل من نسبة التصنيفات الخاطئة"

كما عّرفها بابام Popham بأنها "مقياس لمدى ملائمة أداء الطالب لهدف محدد"

طرق تحديد درجة القطع:

تبنى خبراء القياس والتقويم التربوي العديد من الطرق والمحاولات لتحديد درجة القطع، واختلفت تلك الطرق والأساليب باختلاف خبرات ود رجات تأهيل

وتخصصات هؤلاء الخبراء، فهنالك عدة طرق متعددة لتحديدها والتي ذكر بيرك Berk 1986 أنها بلغت ثمان وثلاثين طريقة مختلفة من حيث أسلوب بناءها وتطبيقها وتفسير نتائجها، ويمكن تصنيفها إلى ثلاثة تصنيفات كالتالي:

أولا: الطرق التحكيمية Judgmental - methods

تستند الطرق التحكيمية على آراء المتخصصين في مجال القياس أو في محتوى الاختبار المراد قياسه، وذلك لتحديد درجة القطع المناسبة، حيث يسترشد بآرائهم في تحديد مستوى الطالب ذو الحد الأدنى من الكفاية (تحديد عدد الأسئلة التي يجب أن يجيب عليها الطالب بصورة صحيحة) واضعين في اعتبارهم المتغيرات التي تؤثر في تحديد هذه الكفاية.

مميزات وعيوب الطرق التحكيمية:

لكل طريقة من الطرق التحكيمية مميزات وعيوب خاصة بها وهنالك عيوب ومميزات تكاد تكون مشتركة بينها وهي كالتالي:

مميزات الطرق التحكيمية:

كما استرسل (بيرك Berk) بأن من مميزات الطرق التحكيمية ما يلي:

1- أنها سهلة الاستخدام مقارنه بالطرق الأخرى.
2- أنها سهلة التطبيق والفهم.
3- أنها سهلة التفسير.
4- لا تحتاج إلى جهد ووقت كبيرين.
5- أنها سهلة الحساب ولا تعتمد على أساليب إحصائية معقدة.

عيوب الطرق التحكيمية:

أشار كل من بيرك، وهامبلتون، وايغنور(Hambleton & Eignor, 1980; Berk, 1986) إلى عيوب استخدام الطرق التحكيمية وهي:

1- اختيارية Arbitrary وسميت كذلك لأن هذه الطرق تعتمد كليا "على آراء المحكمين وتلعب كفاية هؤلاء المحكمون دوراً بارزاً في الوصول إلى المعيار المناسب لتصنيف الطلاب وتحديد بنود الاختبار التي يمكن لأقل الطلاب كفاءة الإجابة عليها بشكل صحيح ويعتقد بأن كون هذه الطرق اختيارية لا يعد سبباً كافياً لرفض هذه الطرق لأن هذه العملية الاختيارية ليست مطلقة ولكنها تعتمد على مجموعة من الأسس في تحديد مستويات الأداء.

2- صعوبة تحديد الطلاب الأقل كفاءة.

3- لا تعتمد على بيانات تجريبية فعلية لكي يسترشد بها المحكمون في تقدير أحكامهم.

4- الحصول على درجات قطع مختلفة عند تطبيقها على اختبار موحد وعينة دراسية واحدة.

5- ذاتية المحكمين في اختيار الحد الأدنى من الكفاية وذلك بسبب اختلاف تصوراتهم حول عملية الإتقان.

6- تفاوت عملية الاتساق الداخلي في عملية التحكيم.

7- صعوبة التنبؤ بأداء الطلاب في الاختبار.

على الرغم من وجود بعض العيوب في استخدام الطرق التحكيمية إلا إنه توجد بعض المعايير المطروحة في أدبيات القياس التربوي التي تحد من تأثير هذه العيوب حيث ذكر بيرك (1986) بعض المعايير نذكر منها ما يلي:

أ- استخدام طرق لتوحيد بيانات الأداء الفعلي.

ب- يجب أن تكون عملية تحكيم مفردات الاختبار متزامنة مع تطبيق الاختبار.

ج- الاهتمام بتدريب المحكمين.

د- تنظيم عملية تحكيم بنود الاختبار.

هـ- إيجاد الحلول المناسبة لتسوية الاختلافات في أحكام المحكمين.

و- تصميم استمارات تنظم عملية التحكيم.

ز- تزويد المحكمين ببيانات مسبقة حول أداء الاختبار.

وتناول كثير من الباحثين الطرق التحكيمية وأسهبوا في ذكر تفصيلاتها وتطبيقاتها ومميزاتها وعيوبها ومن ضمن هؤلاء الباحثين علام 1995، وبيرك 1986، وجايقر 1989 ومن هذه الطرق ما يلي:

1- طريقة إيبل Ebel's method:

اقترح إيبل Ebel طريقة لتحديد درجة القطع تعتمد على تعليقات المحكمين على مفردات الاختبار وذلك من خلال اعتبار أن مواصفات المفردات لها بعدين هما: بعد الصلة Relevance وبعد الصعوبة difficulty ويشتمل بعد الصعوبة على ثلاثة مستويات (سهلة، متوسطة، صعبة) ويشتمل بعد الصلة على أربعة مستويات (أساسية، مهمة، مقبولة، غير مهمة) وبذلك يستخدم المحكم أعداداً مختلفة من تصنيفات الصعوبة والصلة، ثم يطلب من كل محكم ما يلي:

أ- تحديد الخلية المناسبة لكل مفردة من المفردات بناء " على توافق محتواها مع تصنيفات البعدين.

ب- تحديد النسبة المئوية لعدد الطلاب ممن لديهم الحد الأدنى من الكفاية الذين يجب أن يجيبوا عن كل مفردة بصورة صحيحة وذلك لكل خلية من الخلايا، ثم يضرب عدد مفردات كل خلية في النسبة المئوية، ثم يقسم مجموع نواتج جميع الخلايا على عدد المفردات، وبالتالي نحصل على درجة القطع.

عيوبها:

تناول بيرك Berk 1986 عيوب هذه الطريقة حيث ذكر بأن أيبل لم يلتزم بالمواصفات التي تستخدم في كل من البعدين والتمييز بينهما، لكي يسترشد بها المحكمين في عملية التحكيم، لذلك ربما يعتمد كل محكم على مواصفات يحددها لنفسه مما يؤثر بلا شك في دقة التصنيف واتساقه، كما أن هذه الطريقة اعتمدت على

تحكيم المفردة ككل ولم تأخذ بعين الاعتبار بدائل الإجابات التي تشتمل عليها مفردة الاختيار من متعدد. بالإضافة إلى أن طريقة إيبل تحتاج تحكيم خاص يتمثل بنسب مئوية للمفردات وذلك لكل خلية عبر بعدي الصلة والصعوبة المستمرين، فإذا لم تكن هنالك معايير خارجية يبنى عليها هذا الحكم فإنه يبدو إنه سيكون حكماً اعتباطيا.

للحد من هذه العيوب قام كل من سكاكون وكلينق Skakun & Kling بإجراء تعديلات على طريقة إيبل بحيث أصبحت تعتمد هذه الطريقة على بعدين هما : معامل الصعوبة ويشتمل على سهل - متوسط الصعوبة - صعب جد ا (ومعامل التصنيف ويشتمل على واقعي - شامل - حل المشكلات) وتستخدم بيانات الأداء في البنود لتقدير تصنيفات الصعوبة ويحدد كلاً من باني الأسئلة ولجنة الاختبار الأسئلة طبقاً لمستويات التصنيف، بناء. على ذلك تـضرب الأسئلة فـي الخلايا المكونة من المصفوفة 3 × 3 ثم يقوم المحكمين بالإشارة إلى نسبة المفردات في كل خلية والتي يجب أن يجيب عليها الطالب المتقن إجابة صحيحة.

المميزات:

لها نفس مميزات طريقة إيبل ويحدد تصنيف معامل صعوبة البند بالاستعانة ببيانات الأداء، وتضرب في كلا البعدين مما يسهل عمل الحكم.

العيوب:

إن عيوب هذه الطريقة تكمن في عدم وضع عاملي الصلة والأهمية في الاعتبار ومن المحتمل أيضاً أن يصعب على بعد التصنيف تحديد مناطق المحتوى المحددة للحد من هذه العيوب.

قام كل من سكاكون وكلينك Skakun&kling بإجراء تعديلات على طريقة إيبل بحيث أصبحت هذه الطريقة تعتمد على بعدين أيضاً وهما : معاملي الصلة والتصنيف وأصبح معامل الصلة ذو أبعاد ثلاثة (مهمة جدا - مهمة - مقبولة) وهذه الأبعاد مقترنة ببعد معامل التصنيف.

المميزات:

لها نفس مميزات طريقة إيبل وتعتبر عملية ضرب كلا البعدين ميسرة لعمل المحكمين

العيوب:

هذه الطريقة لا تضع في اعتبارها مستويات صعوبة البنود الواقعية، ومن المحتمل أن يصعب في بعد التصنيف تحديد مناطق المحتوى المحددة

2- طريقة أنجوف 1971 Angoffs method

يطلب من كل محكم في هذه الطريقة فحص كل فقرة من فقرات الاختبار، ثم يطلب منهم تصور مجموعة من الأفراد الذين حققوا الحد الأدنى للكفاية التي يقيسها الاختبار، ثم يقدر نسبة عدد الأفراد الذين يحتمل أن يجيبوا إجابة صحيحة عن كل مفردة من مفرداته، ويمثل متوسط هذه النسب الحد الأدنى لمستوى الاجتياز في الاختبار.

مميزاتها:

أ- سهلة التنفيذ.

ب- سهلة الفهم.

ج- أنها تعتمد على تصور عدد من الأفراد الذي حققوا الحد الأدنى للكفاية المطلوبة بدلا من تصور فرد واحد.

عيوبها:

يصعب على كثير من المحكمين تعريف الطلاب من ذوي الحد الأدنى للكفاية، ويصعب أيضاً تقدير القيم الاحتمالية لبعض المفردات المصاغة في عبارات منفية (Negatively) أو المتضمنة عمليات حسابية. (Computational)

للحد من هذه العيوب أشار كل من جايقر وكروس وآخرون ;Jaeger, 1989 Cross et al.، 8419 بأن أنجوف Angoof أجرى تعديلات على طريقته وذلك

لتسهيل عملية تقديرات قيم المحكمين الاحتمالية لأسئلة الاختبار. حيث يطلب من كل محكم تقدير احتمال أن يجيب الطلاب ذوي الحد الأدنى للكفاية كل سؤال دون أن يلجئوا إلى تخمين (Guessing) الإجابة، وتسجل تقديراتهم الاحتمالية في استمارة خاصة تحتوي على مقياس عشري الفترات Equal Interval scale

1= صفر – 0,10	0,20–0,11=2	0,30–0,21=3	0,40–0,31=4
0,50–0,41=5	0,60–0,51=6	0,70–0,61=7	0,80–0,71=8
0,90–0,81=9	1,00–0,91=10		

ثم يطلب من كل محكم اختيار القيمة الاحتمالية التي تمثل تقديره الخاص بكل سؤال من أسئلة الاختبار، ثم تجمع القيمة الاحتمالية التي قدرها كل محكم لكل مفردة، ثم نوجد متوسطها لتمثل درجة القطع.

3- طريقة ندلسكاي Nedlesky's Method

أقترح نيدلسكي Nedelsky 1954 هذه الطريقة لتحديد درجة القطع للاختبارات التي تشتمل على أسئلة ذات اختيار من متعدد حيث تقوم مجموعة من المحكمين المختصين بتحديد درجة القطع المناسبة للاختبار، وفي هذه الطريقة يطلب من كل محكم فحص كل مفردة من مفردات الاختبار، ثم يحدد من بين الخيارات الخاصة لكل مفردة تلك التي يمكن أن يتجنب الطلاب ذوي الحد الأدنى للكفاية اختيارها لأنها لا تمثل في نظرهم الإجابة الصحيحة للمفردة، بعد ذلك يضع المحكم درجة للسؤال. بناء على ما تبقى من الخيارات، وذلك بأخذ مقلوب الخيارات المتبقية.

فمثلاً إذا كانت المفردة تشتمل على خمسة خيارات ورأى المحكم أن الطالب ذو الحد الأدنى للكفاية يمكن أن يتجنب اختيار ثلاثة منها عندها تكون درجة السؤال عبارة عن:

$$0,5\ \frac{1}{2} = \frac{1}{3-5} =$$

بعد ذلك يقوم المحكم بجمع درجات الأسئلة، ويكون المجموع الكلي هو درجة القطع بالنسبة لذلك المحكم، وتجمع في النهاية الدرجات التي يتحصل عليها المحكمون، ثم يحسب المتوسط وبالتالي نحصل على درجة القطع للاختبار ككل، غير أن ندلسكاي وجد إنه من الأفضل إيجاد الانحراف المعياري (σ) لدرجات القطع ثم ضربها في مقدار ثابت (K) يتفق عليه المحكمون وتعتمد قيمة هذا المقدار على النسبة المئوية التي يرى المحكمون أنها مطلوبة لنجاح الطلاب من ذوي الحد الأدنى للكفاية.

ويتم تحديد درجة القطع في هذه الطريقة وفق المعادلة التالية:

$$MPL = \bar{M}\,FD + K\,\sigma\,F\,D$$

$$\bar{M}\,F\,D = \frac{\Sigma\ FD}{N}$$

$$\sigma\,F\,D = \sqrt{\Sigma\ \frac{(MFD - \bar{M}\,FD)^2}{N}}$$

حيث إن:

MPL : تمثل درجة القطع.

MFD: تمثل متوسط الحد الأدنى المحدد من المحكمين.

K: تمثل المقدار الثابت المحدد من المحكمين.

σ F D: تمثل الانحراف المعياري لدرجات القطع

N : تمثل عدد المحكمين.

وأشار بيرك Berk, 1986 إلى بعض عيوب هذه الطريقة وهي كالتالي:

لا يوجد تبرير منطقي يؤيد افتراض (ندلسكاي) أن الطالب الذي لا يعرف الإجابة الصحيحة لسؤال معين يلجأ إلى التخمين العشوائي لأن مشتتات الأسئلة

تصمم عادةً لجذب غير المتمكنين، كما أن هذه الطريقة تسمح فقط بقيم احتمالية منفصلة تعتمد على عدد خيارات الإجابات، لذلك يضطر المحكمون لتعيين القيمة (0,5) لمعظم المفردات، هذا بالإضافة إلى أن هذه الطريقة تحتاج لقدر كبير من التدريب لكي يستخدمها المحكمون استخداماً مناسباً.

4- طريقة انجوف - ندلسكاي المختلطة

Angoff Nedelsky combination, Reod 1984

في هذه الطريقة يتم حساب المعيار على أساس حساب متوسط المعايير الناتجة عن استخدام طريقتي أنجوف وندلسكاي.

المميزات:

لها نفس مميزات طريقة أنجوف.

العيوب:

لها نفس عيوب طريقتي أنجوف وندلسكاي بالإضافة إلى احتمالية زيادة مصادر الخطأ المختلفة وذلك عندما نحسب متوسط المعيارين.

5- طريقة تقدير أهمية الصعوبة difficulty importance estimate

في هذه الطريقة يقوم المحكمون بفحص كل هدف، وذلك لتقدير الحد الأدنى للنجاح ويجب أن تتفق عينة المحكمين على الحد الأدنى وعلى كل هدف أيضاً، وذلك لأن للحد الأدنى وزن خاص طبقاً لأهمية صلته بالهدف المنشود، ثم تضرب النسبة المئوية لدرجة النجاح في نسبة أهمية الدرجة لكل هدف، والناتج يجمع من خلال جميع الأهداف، ويضرب في عدد بنود الاختبار وذلك للوصول للمعيار، ويجب أن يضع الحكام في الاعتبار الخطأ المعياري وذلك عند اتخاذهم القرارات الخاصة بدرجة القطع.

المميزات:

أ- سهلة الفهم.

ب- سهلة الاستخدام وتعمل على ربط درجة القطع بقائمة خصائص النطاق السلوكي.

العيوب:

استخدمت تقديرات المحكمين للأهداف لتقدير معيار بنود الاختبار، بينما نجد إنه من المفروض أن تكون العلاقة بين الأسئلة والأهداف مباشر ة، وليس بالضرورة افتراض أن تقديرات الأسئلة تساوي تقديرات الأهداف، ومن المحتمل أن يكون نوع الخطأ المعياري المستخدم غير ملائم لتفسير درجة القطع.

6- طريقة مواصفات المفردة (ميلز و بار 1983م):

في هذه الطريقة يقوم المحكمون بقراءة مواصفات كل مفردة والتي تحتوي على المهارة المراد قياسها، ووصف محتوى البند المقبول، وخواص المثير، والقوانين التي تحكم إحداث خيارات الاستجابة وعينة المفردات، ومن ثم فإن على المحكمين إدراك وتصور الإطار الكلي للمفردات والمتمثل في المواصفات والتنبؤ بنسبة البنود في الإطار الكلي، والتي يمكن أن يجيب عليها الممتحن ذو الحد الأدنى من الكفاية إجابة صحيحة.

المميزات:

سهلة الفهم، وسهلة الاستخدام، وتقوم بربط درجة القطع بمواصفات النطاق السلوكي، وربطهما بكافة البنود التي تقيس النطاق السلوكي، وتستغرق زمناً أقل مما لو عمدنا إلى تقدير كل مفردة بشكل منفصل، وتكون درجة القطع فيها مرتبطة بالنطاق السلوكي وليس بصيغة الاختيار المحدد.

العيوب:

تستخدم تقديرات المحكمين لمواصفات البنود لتقدير درجة القطع لمفردات الاختبار، بينما العلاقة بين المفردات ومواصفات البنود مباشرة، لذا ليس من الضروري أن تكون تقديرات درجة القطع مماثلة لتقديرات مواصفات البنود.

ثانيا: الطرق التي تعتمد جزئيا على التحكيم وتسترشد ببيانات تجريبية:

أشار علام 1995م، بأن الباحثين أدركوا قصورا ملحوظا في الطرق التحكيمية ومن ذلك القصور اعتمادها على تحكيم مفردات الاختبار دون النظر إلى الأداء الفعلي للمختبرين مما يؤدي ذلك إلى الحصول على معايير غير متميزة بدرجة عالية من الصدق.

وأشار كل من سكون وجاليون وفيرار ا بأن درجة القطع المحددة باستخدام بعض من الطرق التحكيمية تؤدي إلى رسوب أكثر من نصف المتقدمين للحصول على ترخيص مهنة الطب في أمريكا، أما في حالة استخدام الطرق التجريبية فإنه يتبين خطأ تلك الدرجات حيث يجتاز معظم الراسبين الحد الأدنى لمستوى النجاح.

لذا اقترح علماء القياس العديد من الطرق المختلفة لتحديد مستويات الأداء التي تتضمن طرقاً تجريبية محاولين في ذلك تجاوز عيوب الطرق التحكيمية والحصول على مستويات أداء دقيقة يمكن الاعتماد عليها ومن هذه الطرق ما يلي:

1- الطريقة التحكيمية المعززة بالمعلومات Informed Judgment method (popham, 1981)

يتضح من مسمي هذه الطريقة أنها تعتمد على بيانات مسبقة عن أداء الطلاب حيث تفيد هذه البيانات في إعطاء دلائل للمحكمين عن المستوى العام للطلاب، وبالتالي توضع هذه البيانات في اعتبار المحكمين لكي يسترشدوا بها أثناء تحديدهم للمعايير وفي هذا الشأن ذكر بيرك (1986) بأن هذه الطريقة تعتمد على بيانات

تتعلق بالأداء الفعلي للأفراد، وتتضمن هذه البيانات إحصاءات خاصة بالمفردات مثل معاملات الصعوبة والتمييز ومتوسط درجات الاختبار بعد تطبيقه على العينة المناسبة من الأفراد، وبعض تعليقات وأحكام الخبراء حول مفردات الاختبار، واقتراحات المختصين مثل المدرسين والإداريين وأساتذة الكليات وأعضاء مجالس المدارس، وتقدم جميع هذه المعلومات للمحكمين لفحصها وتبادل المقترحات حولها ومناقشتها وإعادة النظر فيها وذلك لكي يتم التوصل إلى اتفاق حول درجة القطع المناسبة.

المميزات:

سهلة التطبيق وسهلة الفهم ويسهل حسابها، كما أنها تجمع وتوحد بيانات الأداء والبيانات المتميزة من مدى واسع بواسطة المجموعات المختصة بتلك البيانات.

العيوب:

من عيوب هذه الطريقة أنها تتطلب جمع كثير من البيانات التجريبية والمعلومات المتعلقة بالاختبار ومحتواه من مصادر متعددة، وقد يستغرق ذلك وقتاً وجهداً كبيرين، ويبدو أن تحليلات المحكمين للبيانات غير منظمة، كما تبدو الحاجة لوجود خطوط تفصيلية تسهل استخدام أنواع البيانات المختلفة والمناقشة الجماعية للأداء الفردي، والتي لا تبدو مفيدة في هذه الطريقة بسبب تأثيراتها المعيارية.

2- طريقة أنجوف التجريبية المعدلة Modified Angoff empirical

ذكر بيرك أن هذه الطريقة عبارة عن تعديل لطريقة أنجوف التحكيمية، وتتطلب هذه الطريقة المعدلة تزويد المحكمين ببعض البيانات المستمدة من التجريب الميداني لمفردات الاختبار، وذلك لكي يسترشد بها المحكمين في عملية التحكيم، ويطلب من كل محكم تقدير احتمال إجابة الأفراد ذوي الحد الأدنى من الكفاية المطلوبة لمفردات الاختبار ككل وليس من كل مفردة على حده كما في الطريقة السابقة (التحكيمية)،

وبعد الحصول على هذه التقديرات يتم تزويد المحكمين ببيانات تجريبية عن مفردات الاختبار تشتمل على متوسط صعوبة كل مفردة، ثم يطلب منهم تعديل تقديراتهم الاحتمالية في ضوء تلك البيانات التجريبية، ويكون الحد الأدنى لمستوى الاجتياز في الاختبار هو متوسط مجموع تقديراتهم المعدلة.

المميزات:

سهلة التطبيق، سهلة الفهم ويسهل حسابها وتعمل على ربط المعيار بالكفايات المقاسة، وتدمج معامل صعوبة البند الحقيقي خلال عمليات التحكيم.

العيوب:

إن عملية استخدام تقديرات المحكمين للكفايات يجب أن تكون على درجة عالية من التوافق والانسجام، وذلك بين المفردات والكفاية ولا يمكن افتراض أن تكون التقديرات الواقعية متطابقة. وأضاف جايقر Jaeger بأنه أجريت تعديلات في عام 1978م على طريقة أنجوف الأساسية وذلك بهدف تبسيط مهمة المحكمين بحيث يطلب من كل محكم تقييم كل مفردة من مفردات الاختبار بالإجابة (بنعم) أو (لا)، ومن ثم يتم تحديد الحد الأدنى للكفاية وبعد الحصول على هذه التقديرات يستعرض عدد مختلف من المحكمين ثلاث أنواع من البيانات التجريبية، وذلك خلال ثلاث لقاءات منفصلة وتشتمل هذه البيانات التجريبية على التوزيع التكراري لدرجات الطلاب في الاختبار والأساليب الإحصائية الوصفية ودرجات القطع المحددة من قبل المحكمين، ثم يطلب من عينات المحكمين المختلفة تعديل تقديراتهم بناء. على تلك البيانات التجريبية، وتصبح درجة القطع عبارة عن متوسط تلك التقديرات التي حددتها عينات المحكمين في اللقاءات الثلاثة المنفصلة.

المميزات:

تتيح للمحكمين فرصة لتحسين تقديراتهم بناء. على ثلاث أنواع من البيانات

المختلفة، ويحدد درجة القطع عينات مختلفة من المحكمين وتميل هذه الطريقة إلى زيادة ثبات درجة القطع.

العيوب:

تحديد الإجابة بنعم أو لا، بالإضافة إلى أن هذه الطريقة تفرض على المحكمين أن يلتقوا ثلاثة لقاءات منفصلة، كما أنها تستغرق كثيراً من الوقت، وتعتبر أكثر تعقيداً من الطرق السابقة، ويعتبر تفسير المعيار النهائي صعب جداً بالنسبة للمحكمين.

لتلافي بعض من هذه العيوب قام كل من سنديرس ومابوس بإجراء تعديلات على الطريقة فأصبحت تحتوى على عينة واحدة من المحكمين، وأصبح تحديد المعيار النهائي يتم خلال لقاء واحد يجمع بين المحكمين.

كما قام شيبرد أيضاً بإجراء تعديلات على طريقة أنجوف إلا ساسية بحيث أصبحت هذه الطريقة تتطلب أربعة أنواع من المعطيات التي يجب أن توضع بعين الاعتبار وهي:

أ- أحكام مطلقة مبنية على فحص المفردات (مثل طريقة أنجوف).

ب- بيانات تجريبية على عينة محكمة صنفت المتقنين وغير المتقنين (مثل طريقة المجموعات المتضادة).

ج- تقديرات صانعي القرار (المحكمين) لمستوى النجاح المقبول.

د- التوفيق بين تقديرات المحكمين المختلفة.

فإذا كانت درجات القطع الممكنة والمختلفة قد نشأت من هذه المعطيات فيجب على محدد المعايير أن يوفق بين هذه الاختلافات.

المميزات:

من مميزاتها أنها تضم أربعة أنواع من البيانات التجريبية والتحكيمية وتضع في حسابها معظم المصادر الرئيسة للقيام بعملية تحديد المعايير، وتحتفظ بمميزات طريقتي أنجوف والمجموعات المتضادة.

العيوب:

لها نفس عيوب طريقتي أنجوف والمجموعات المتضادة فضلا عن صعوبة التوفيق بين الاختلافات بين المعايير وتعديل المعايير المبنية على معدلات النجاح المقبولة وعدم الثبات، بينما نجد أن إستراتيجية (التثليث) لها مميزات معينة وخطوط إرشادية محددة، والتي يمكن أن تكون مقبولة بشكل كبير حتى تصل بشكل منظم إلى المعيار النهائي.

3- طريقة توفق بين الطرق المطلقة والطرق النسبية Absolute Relative Compromise (I)(II)

أشار بيرك إلى أن هذه الطريقة تعمل على التوفيق بين أحكام المحكمين المتعلقة بأداء الأفراد في الامتحان والأداء الفعلي المستمد من البيانات التجريبية المبنية من تطبيق الاختبار على العينة، وفي هذه الطريقة يطلب من المحكمين تحديد قيمتين هما:

أ - الحد الأدنى للنسبة المئوية لعدد المفردات التي يجب أن يجيب عليها الفرد إجابة صحيحة لكي يعتبر ناجحاً في الاختبار، ويرمز لهذه النسبة بالرمز (K).

ب- النسبة المئوية لعدد الأفراد المتوقع اجتيازهم للاختبار ويرمز لهذه النسبة بالرمز (V)، ويمكن التوصل إلى معادلة تربط بين النسبتين K، V، والتي يمكن استخدامها لتعديل تلك القيم التي لا تتفق مع التوزيع الفعلي لدرجات الاختبار، إذ يمكن رسم منحنى للمتوسطات والانحرافات المعيارية لقيم K، V التي حددها المحكمون، وكذلك النسبة المئوية لعدد الأفراد المتوقع نجاحهم ويرمز لها بالرمز Y كدالة تناقصية لدرجات الاختبار ويرمز لها بالرمز X، ثم يرسم منحنى لتحديد النقطة ويرمز لها بالرمز (I) والتي تمثل الحد الأدنى لمستوى النجاح، بحيث توفق بشكل أمثل بين المستويات المطلقة (التحكيم) والمستويات النسبية (الأداء الفعلي للأفراد في الاختبار).

المميزات:

من مميزات هذه الطريقة أنها توفق بين الأحكام المتعلقة بمفردات الاختبار وبين الأداء الفعلي للمختبرين أي بين المستويات النسبية.

العيوب:

من عيوبها أنها تتطلب من المحكمين أن يوفروا تقديرين وهذا يجعل تنفيذ هذه الطريقة أكثر صعوبة من غيرها سواء في عملية التحكيم أو من الناحية الإحصائية، وكذلك تتطلب جهداً كبيراً من جانب المحكمين، بالإضافة إلى أن تفسير المعيار من قبل المتخصصين قد يكون صعب جداً.

لذلك أجرى هوفستي تعديلات لهذه الطريقة حيث يتطلب من المحكمين تحديد أربع قيم هي:

أ- أعلى نسبة مئوية لعدد المفردات التي يمكن أن يجيب عليها الفرد إجابة صحيحة ويرمز لها بالرمز (Kmax) وهي عبارة عن درجة القطع المرضية حتى إذا حصل عليها جميع الطلاب.

ب- أدنى نسبة مئوية من عدد المفردات التي يمكن أن يجيب عليها الفرد إجابة صحيحة ويرمز لها بالرمز (Kmin) وهي عبارة عن درجة القطع الدنيا التي لا يمكن لأي الطالب أن يحرز أقل منها.

ج- أعلى نسبة مئوية مقبولة من الراسبين ويرمز لها بالرمز (Fmax)

د- أدنى نسبة مئوية مقبولة من الراسبين ويرمز لها بالرمز (Fmin)

وباستخدام توزيع درجات الاختبار المتجمع (التراكمي) فإننا نستطيع تحديد العلاقة التجريبية بين (K و F)، ويكون المعيار هو نقطة التقاطع بين النموذج (المطلق) والمنحنى (النسبي).

المميزات:

من مميزات هذه الطريقة أنها توفق بين الأحكام المتعلقة بمفردات الاختبار وبين الأداء الفعلي للمختبرين، أي بين المستويات المطلقة والمستويات النسبية كما تضع في اعتبارها الحد الأعلى للنسبة المئوية للمفردات التي من الممكن أن يجيب عليها الفرد إجابة صحيحة (Kmax)، كما تقوم بتبديل النسبة المئوية من الأفراد المتوقع اجتيازهم للاختبار) (V) بأعلى نسبة مئوية مقبولة من الراسبين Fmax وبأدنى نسبة مئوية مقبولة من الراسبين (Fmin).

العيوب:

من عيوب هذه الطريقة أنها تتطلب جهداً كبيراً من المحكمين، وتتطلب توفير أربعة تقديرات وهذا يجعل تنفيذ هذه الطريقة أكثر صعوبة من غيرها سواء. في عملية التحكيم أو الإحصاء.

ثالثا: طرق تعتمد على البيانات التجريبية وتسترشد بالتحكيم Empirical Judgmental

1- طريقة المجموعات المحكية CRITERION GROUPS

أشار كل من بيرك وهامبلتون وآخرون (BERK. 1986; Hambleton et al,1978) إلى أن هذه الطريقة تتطلب انتقاء مجموعتين من الطلاب الأولى (متقنة) تلقت التعليم (والثانية غير متقنة) لم تتلق التعليم، ويكون مدى توزيع درجات الطلاب من الدرجة (صفر) إلى الدرجة (ن) حيث (ن) تمثل عدد الأسئلة التي تقيس كل هدف على حده، وتتطلب الطريقة تقسيم الطلاب بناء اً على درجة القطع (س) إلى قسمين (متقنين – غير متقنين)، وبافتراض أن الطلاب في المجموعة المتقنة (التي تلقت التعليم) عبارة عن طلاب متقنين فإنه يمكن تقسيمهم بناء. على درجة القطع إلى فئتين:

الأولى: متقنين بالفعل TM

والثانية: غير متقنين غير حقيقين (FN).

وبالمثل يتم تصنيف طلاب المجموعة الثانية غير المتقنة (التي لم تتلق التعليم) إلى فئتين الأولى غير متقنين فعلاً(TN)) والثانية متقنين غير حقيقين. (FM)

وتعتبر توزيعات درجات الطلاب في المجموعتين المتقنة وغير المتقنة المحددات الأولية (Primary Determinants) لمدى دقة تصنيف الاختبار للطلاب كمتقنين وغير متقنين فعلاً للهدف المراد قياسه.

وتحدد قيمة درجة القطع (س) من خلال فحص خلية الاحتمالات لكل درجة قطع محتملة، وتكون درجة القطع المثلى تلك التي تزيد نسبة التصنيفات الصحيحة (TM – TN) وفي المقابل تحجم نسبة التصنيفات الخاطئة. (FN – FM)

ويتم تحديد درجة الإتقان بشكل أولي من خلال توزيعات تكرارات درجات الطلاب في المجموعتين المتقنة (التي تلقت التعليم) وغير المتقنة (التي لم تتلق التعليم) حيث يصنف الطلاب بناء على نقطة تقاطع التوزيعين إلى فئتين متقنين بالفعل وغير متقنين بالفعل للهدف المقاس.

ثم استخدم بيرك (1976) Berk طريقة بديلة تهدف لإيجاد معامل الصدق لكل درجة قطع محتملة (س)، لذا فإن متغيرات التصنيف تعطي قيم ثنائية هما:

أ- متغير التصنيف المنبئ The Predictor Classification

حيث يقدر للطالب القيمة (1) عندما تكون درجته أعلى من درجة القطع، ويقدر للطالب القيمة (صفر) عندما تكون درجته أدنى من درجة القطع.

ب- متغير التصنيف المحك Criterion Classification

يقدر لكل طالب ضمن مجموعة المتقنين القيمة (1) بينما يقدر لكل طالب ضمن مجموعة غير المتقنين (صفر)، ويحسب معامل الصدق من قيمة معامل (فاي) بين المتغيرين الثنائيين.

يلي ذلك حساب المنفعة والضرر النسبيين لأخطاء التصنيف، ويمكن تمثيل قيمة الضرر للخطأ (FN) بالقيمة (-1) وللخطأ (FM) بالقيمة (-2)، في المقابل فإنه يتم حساب الفائدة النسبية لتحديد الإتقان بالفعل (TM) بالقيمة (2) ولتحديد عدم الإتقان بالفعل (TN) بالقيمة (1).

مميزاتها:

أنها نموذج للصدق التكويني وتعتبر هذه الطريقة ذات حساسية خاصة للخصائص الإحصائية للمفردة وكذلك الصدق البنائي.

عيوبها:

من عيوبها أنها لا تحدد المعيار الحقيقي وفي أفضل الحالات تقوم بتحديد موقع الأداء على متصل الدرجات حيث يتعذر التمييز بين المجموعتين المتقنة وغير المتقنة.

2- طريقة المجموعات المتضادة (LIVINGSTON& ZIEKY,1982)
CONTRASTING GROUPS

ذكر كل من بيرك وجايقر بأن هذه الطريقة مماثلة لحد ما لطريقة المجموعات المحكية وطريقة المجموعة الحدية، ولكنها تختلف في كيفية تحديد درجة القطع في الاختبار، ففي هذه الطريقة يقوم المحكمون المتخصصون بتحديد مجموعتين من الطلاب بحيث تكون الأولى متقنة والأخرى غير متقنة بناء. على ما يلي:

المعلومات المتوفرة حول مستويات مهارات الطلاب.

استناداً على درجات الطلاب وأداءهم داخل الفصل حيث يتم تصنيف الطلاب إلى فئتين الأولى الحاصلون على تقدير (جيد جداً فما فوق) والثانية الطلاب الحاصلون على تقدير (مقبول فمادون).

ثم يطبق الاختبار على كلتا المجموعتين ويرسم منحنيان يمثلان التوزيعين التكراريين لدرجات الطلاب في الاختبار، وتمثل نقطة تقاطع المنحنين درجة قطع الاختبار،

هذا الاقتراح يفترض تداخل التوزيعين التكراريين وعدم تطابقهما، ويمكن تعديل هذه الدرجة آخذةً في اعتبارها خطأي التصنيفين، أي يمكن اختيار الدرجة التي تقلل من خطأ التصنيف الموجب POSITIVE FALSE أو خطأ التصنيف السالب. (FALSE NEGATIVE).

عيوبها:

لها نفس عيوب طريقة المجموعات المحكية، بالإضافة إلى إنه من الصعب تجاوز الخطأ الناجم من المحك.

3- طريقة المجموعة الحدية BORDERLINE GROUP (LIVINGSTON & ZIEKY, 1982)

ذكر كل من بيرك وجايقر (BERK, 1986; JAEGER, 1989) بأن هذه الطريقة تتطلب من المحكمين انتقاء عينة من الطلاب الذين يرون بأن كفايتهم المراد قياسها تقع عند الحد الفاصل بين التمكن وعدم التمكن، ومن المحتمل أن تكون نسبة الطلاب قليلة بالنسبة للمجموعة الكلية، ثم يطبق الاختبار على أفراد العينة وبعد ذلك يتم إيجاد وسيط درجات هذه المجموعة ليكون بمثابة مستوى الأداء المطلوب.

مميزاتها:

تمتاز هذه الطريقة بأنها سهلة الفهم والتطبيق والإحصاء غير أنها تحتاج إلى عدد كاف من الطلاب وذلك لكي نحصل على تقديرات مستقرة نسبياً.

عيوبها:

من عيوبها إنه يصعب على المحكمين أو المدرسين انتقاء المجموعة الحدية، إذ أن تعريف أو تحديد الطالب الذي يقع عند الحد الفاصل بين الإتقان وعدم الإتقان أصعب بكثير من تحديد الفرد المتقن أو غير المتقن، كما تتدخل بعض المتغيرات في عملية انتقاء أفراد هذه المجموعة وقد يصعب تحديد الحجم المناسب للعينة.

4- طريقة المحك معياري المرجع NORM REFENCEED CRITERION

ذكر بيرك (1986) BERK بأنه في هذه الطريقة يتم اختيار المعيار بناءً على العلاقة بين درجات القطع في الاختبار محكي المرجع وبين درجات الأداء في الاختبار معياري المرجع (NRT) ويتم اختيار درجة القطع أولاً في الاختبار معياري المرجع بطريقة اختيارية على سبيل المثال: مناظرة الدرجة الخام بالدرجة المئينية 50أو75 ثم تختار درجة القطع في الاختبار محكي المرجع على أساس أنها تزيد ما يلى:

أ - نسبة ثبات القرارات الناتجة عن تطبيق الاختبار معياري المرجع ويرمز لها بالرمز (P)

ب- نسبة ثبات القرارات المصححة من عامل الصدفة ويرمز لها بالرمز (K) .

ج- تعيين دالة الفقد Loss Function ويرمز له بالرمز (L).

المميزات:

من مميزاتها أنها سهلة التطبيق والفهم والإحصاء ويمكن توفير بيانات محك الاختبار معياري المرجع بسهولة ويسر، بالإضافة إلى أن هذه الطريقة تضع في اعتبارها دالة الفقد(LOSS FUNCTION) والعلاقة المباشرة بين الاختبارين.

العيوب:

إن استخدام درجات القطع الاختيارية في الاختبار معياري المرجع كمحك يعد خطأ فني وغير منطقي.

5- طريقة النتائج التربوية EDUCATIONAL CONSEQUENCES (BLOCK, 1972)

ذكر بيرك (1986) BERK بأن هذه الطريقة تعتمد أساسا على اختيار المعيار من خلال العلاقة بين درجات الاختبار ومحك التعلم المستقبلي نتائج انفعالية أو

معرفية حيث يتوقع أن يزداد المنحنى طردياً ليربط بين المتغيرين، وتحدد درجة القطع وفقاً على أثرها في زيادة أداء المحك.

المميزات:

سهلة التطبيق والفهم والحساب.

العيوب:

إن معيار (التعلم المستقبلي) يحتاج إلى تحديد واضح ولا تبدو الافتراضات في هذا النموذج متطابقة مع البيانات الواقعية، بالإضافة إلى أن العلاقة بين المتغيرات المقترحة لا توفر "دالة درجية" (STEP·FUNCTION) أو خاصية فاعلة في المحك لتحديد درجة القطع.

6- طريقة إيمريك لتقييم اختبار التمكن

EMRICK'S MASTERY TESTING EVALUATION

ذكر إيمريك (Emrick, 1971) بأن هذه الطريقة تناسب الاختبارات الفرعية القصيرة التي تتكون من خمس فقرات أو أقل، وتعتمد هذه الطريقة على نظرية القرار، إذ يتم من خلالها تحليل ثلاثة من أخطاء القرار وهي الخطأ الأول (α) والخطأ الثاني (β) ونسبة الفاقد لأخطاء

القرارين الأول والثاني (RR) لذلك يجب تحديد الخطأ الأول (α) والخطأ الثاني (β) ونسبة الفاقد وطول الاختبار، ويتم حساب نسبة الفاقد من خلال تقييم الأنواع المختلفة لأخطاء القرار، ومن ثم تسجيل مجموع المخاطرات بتحديد درجة القطع باستخدام المعادلة التالية:

$$\text{Log}\,(\beta\, K = \;/\; -1\ \alpha) + 1/n(\text{Log RR}) - \text{Log}\,(\alpha\,\beta)\ \ /(-\,\alpha)\,(1-\beta 1)$$

حيث إن:

k: ترمز إلى درجة القطع

α : ترمز إلى الخطأ الأول

β : ترمز إلى الخطأ الثاني

RR: ترمز إلى نسبة الفاقد للخطأين الأول والثاني.

N: ترمز إلى عدد الأسئلة.

7- طريقة نوفيك ومعاونيه The Work Of Novick & Collaborators

ذكر (MESKAUSKAS 1976) بأن هذه الطريقة عبارة عن مزيج من الطرق المختلفة، وتهدف إلى تحديد كل من درجة القطع والتوزيع الاحتمالي للأداء والفواقد الناتجة عن الخطأ الأول (α) والخطأ الثاني (β) ويفيد تحديد هذه المتغيرات في زيادة دقة القرارات المتخذة، حيث، يتم وضعها في شكل توزيع احتمالي، ويضاف إلى هذا التوزيع درجات الاختبار بحيث ينتج عنه توزيع احتمالي أكثر حساسية، ويتم تحديد القيمة المقدرة للمستوى الحقيقي للطالب المتقن عن طريق تحديد عينات من المفردات، ثم يضرب الاحتمال الذي تزيد فيه الدرجة الملاحظة على التحصيل المعطى، ثم تقارن النتيجة بالدالة المشابهة التي تشير إلى الاحتمال الذي لا تزيد فيه الدرجة الملاحظة.

8- طريقة رودابش للدرجة الحقيقية الثنائية

ROUDABUSH'S DICOTOMOUS TRUE SCORE

ذكر (1976) MESKAUSKASبأن طريقة رودابش تستخدم مع الاختبارات ذات الفقرة الواحدة، وتشتمل هذه الطريقة على نموذجين من النماذج القطعية وكلا النموذجين يشتملان على مقياس ثنائي للدرجة الحقيقية، وتقوم هذه الطريقة على تحديد أربعة من الأخطاء المحتملة وهي :

(α 1 α2) احتمال أن يكون الفرد المتقن ضمن غير المتقن في المحك و(β2 β1) احتمال أن يكون الفرد المتقن ضمن غير المتقن في المحك. ويتم تحديد الأخطاء الأربعة وفق العلاقات التالية:

$(\alpha 2 = P (x$ [illegible] $Xc / T = 0) \alpha 1 = P (x$ [illegible] $Xc / T = 0$

$\beta 2 = P (x$ [illegible] $Xc / T = 0) \beta 1 = P (x$ [illegible] $Xc / T = 0)$

حيث إن:

T: الدرجة الحقيقية

X: الدرجة الناتجة

Xc: درجة القطع

وتستخدم العلاقات التالية لتحديد درجة القطع وهي كالتالي:

$F 00 = N0 (1 - \alpha 1) (1 - \alpha 2) + N1 \beta 1 \beta 2$

$(F 01 = N0 (1 - \alpha 1) \alpha 2 + N1 \beta 1 (1 - \beta 2$

$(F 10 = N0 \alpha 1 (1 - \alpha 2) + N1 \beta 2 (1 - \beta 1$

$(F 11 = N0 \alpha 1 \alpha 2 + N1 (1 - \beta 2) (1 - \beta 1$

9- طريقة ملمان ذات القرار المبني على الحدين

MILLMAN'S BINOMIAL BASED DECISION

ذكر كل من هامبلتون، وآخرون وميسكاوسكاس (Hambleton et al, 1978) بأن ملمان طور مجموعة من الجداول التي تعتمد على التوزيع ذو الحدين، والذي يشير إلى كل من تقديرات الخطأ المتوقعة لمجموعة متعددة من الدرجات الحقيقية وطول الاختبار ودرجة القطع، وهذه الطريقة تقوم على افتراض أن الاختبار يتكون من مجموعة عشوائية من الدرجات تتراوح مابين (1- صفر)، ويستخدم التوزيع ذو الحدين لإيجاد البيانات حول قيم محددة مفترضة للدرجة الحقيقية في كل من أطوال الاختبار والاختبارات الكبيرة والتوزيع الهندسي الفوقي أو ذي الحدين، فيتم التنبؤ بالتكرارات المتصلة، ثم يرسم التكرار المتصل للنسب المئوية المختلفة من الأسئلة التي تمت الإجابة عليها بشكل صحيح.

10- طريقة ديفز ودايموند
THE DAVIS DIAMOND BAYESIAN Method

أشار ميسكاوسكس MESKAUSKAS 1976 إلى أن هذه الطريقة تعتمد على نموذج بييز، وقد حدد كل من ديفز ودايموند جدولاً يضع تقديرات تعتمد على الدرجة الملاحظة واحتمال أن يكون المستوى الحقيقي لكفاءة الطالب بالقرب أو أعلى من المستويات المختارة.

بالنسبة للاختبار (ذو خمس فقرات) فإن احتمال أن يكون الطالب قد أحرز خمس درجات وله مستوى كفاءة حقيقي يساوي (0,99) أو أعلى من ذلك يساوي (0,1135) وهو رقم أدنى، ويعرض الجدول درجات القطع الضرورية في مستويين من التصنيفات الحقيقية هما (0,85) ودرجة القطع للاختبارات حتى عشرون فقرة. فإذا أراد صانع القرار أن يقتطع نسبة 85٪ من التصنيفات الحقيقية للطلاب مع وجود مستويات كفاية عند نسبة 90٪ أو أعلى فيجب أن تكون هنالك مجموعة متساوية لعدد المفردات في الاختبار، ولتحقيق هذا الهدف يجب استخدام اختبار تتراوح فقراته مابين 12 إلى 20 فقرة.

أسس تقويم طرق تحديد درجة القطع:

ذكر بيرك (Berk, 1986) بأنه يوجد معيارين لتقويم طرق تحديد درجة القطع وهما كما يلي:

1- الكفاية الفنية Technical - Adequacy

أ - أن تحدث الطريقة تصنيفاً مناسبا للمعلومات: بمعنى أن تحدث الطريقة تصنيفات ثنائية مثل الإتقان / عدم الإتقان، الفشل / النجاح، ويعتبر التصنيف أساسياً لمعظم القرارات المتخذة.

ب- أن تكون حساسة لأداء المختبر: يجب أن تكون الطريقة حساسة لمستويات صعوبة الأسئلة المختلفة ولأي خصائص أخرى مثل معامل التمييز.

ج- أن تكون الطريقة حساسة لكل من المتعلم أو لعملية التدريب: يجب أن تراعي الطريقة عملية التعلم والتدريب الفعلي الذي تلقاه الطالب بحيث تقاس المهارات والكفايات التي تلقاها الطالب بالفعل.

د- أن تتضمن الطريقة طرقاً إحصائية مناسبة: يجب استخدام الطرق الإحصائية المناسبة، ويجب أن تفسر بشكل صحيح.

هـ- أن تحدد الطريقة المعيار الصحيح: بمعنى إنه يجب أن تراعى الطريقة أخطاء القياس.

و- أن تثبت الطريقة صدق القرار: بمعنى إنه يجب الحصول من الطريقة تقدير احتمالات قرارات التصنيفات الصحيحة والخاطئة.

2- الجانب التطبيقي (العملي) Practicability

أ - أن تكون الطريقة سهلة التطبيق Easy to implement : يجب أن تكون الخطوات المستخدمة في تطبيق الطريقة واضحة وسهلة الفهم للمشاركين فيها، وأن تتم في قدر معقول من الوقت.

ب- أن تحسب الطريقة بسهولة Easy to Compute : يجب أن تكون الطرق الإحصائية المستخدمة للحصول على المعيار النهائي سهلة الاستخدام سواء باستخدام الآلة الحاسبة أو البرامج الإحصائية المتوفرة في الحاسب الآلي الشخصي أو المركزي.

ج- أن تكون قابلة للتفسير بسهولة للعامةEasy to interpet to lay people : يجب أن تكون الطريقة المستخدمة في تحديد درجة القطع قابلة للتفسير وقابلة للفهم، ويجب أن تكون تفسيرات الطريقة واضحة وسهلة الإدراك والفهم للعامة والمهتمين على حد سواء.

د- أن تكون الطريقة جديرة بثقة العامة Credible to lay people: يجب أن تكون الطريقة مقنعة وجديرة بثقة العامة.

كما أضاف بيرك (Berk, 1982) بأن هناك عدة عوامل تساعد على اختيار الطريقة المناسبة وهي:

هـ- مدى أهمية القرارات المبنية على تحديد درجة القطع.

و- الكمية المتاحة من الوقت لتحديد درجة القطع.

ز- مدى توفر المصادر (المادية - الطاقة البشرية) لإنجاز مهمة تحديد درجة القطع.

ح- كفاية المحكمين) بعض الطرق تتطلب دراية أكثر بالمحتوى المراد قياسه وبمستوى الطلاب التحصيلي.

ط- مدى مناسبة طريقة تحديد درجة القطع لنوعية الاختبار المراد قياسه.

الاختبارات التشخيصية

ومن أنواع الاختبارات المحكية الاختبارات التشخيصية التي تهدف للتحقق من اكتساب المتعلم كفايات أو مهارات أساسية تعبر عن نواتج تعليمية محدودة ومحددة، وتشخيص الصعوبات التي تصادفه أثناء تعلمه أو تدربه، والتعرف على مصادر الأخطاء سواء كانت ناجمة عن سو الفهم أو عن عدم التمكن من الإجراءات أو العمليات التي تنطوي عليها هذه الكفايات أو المهارات. وبذلك تساعد المعلم في تصميم أساليب تعليمية علاجية مناسبة تيسر على المتعلم تصحيح هذه الأخطاء ومتابعة التعلم من أجل تحقيق الكفايات أو المهارات المرجوة

خطوات بناء الاختبارات التشخيصية :

الخطوة الأولى: تحديد الكفايات أو المهارات الرئيسة المرجوة ومحتواها:

فكل برنامج تعليمي أو تدريبي يهدف لتحقيق مجموعة من الكفايات أو المهارات

الرئيسة لدى المتعلمين. ونقصد بالكفاية مجموعة متكاملة من المعارف والمهارات الوظيفية المحددة تحديداً دقيقاً والمتعلقة بمجال تعليمي أو تدريبي معين بحيث يمكن تحقيقها وقياسها من خلال البرنامج. ويمكن إجراء ذلك عن طريق:

1- الاستعانة بمجموعة من خبراء المادة الدراسية الذين يقومون بتحديد الكفايات أو المهارات الرئيسة التي يرون أهمية تحققها لدى المتعلمين.

2- التحليل المتعمق لمحتوى المنهج الدراسي أو المجال التدريبي.

3- إجراء دراسة لتقدير احتياجات الفئة المستهدفة.

4- في المجال التدريبي يمكن التوصل إلى المهارات والكفايات اللازم إتقانها في مهنة معينة عن طريق تحليل العمل.

وتتطلب تحديد الكفايات اعتبارات يمكن إيجازها:

1- مدى اتساع الكفاية.

2- إمكانية تعليم الكفاية.

3- قابلية انتقال الأثر.

4- تمثيل الكفايات للسلوك الختامي المستهدف.

إن تحديد الكفايات أو المهارات الرئيسة في الخطوة الأولى ليس كافيا لبناء الاختبار التشخيصي. فالمهارات الرئيسة تعد بمثابة نواتج مركبة وتتضمن معارف وعمليات عقلية ونفسية حركية. لذلك فهي تتطلب تحليلا يرتب مكوناتها ترتيبا منطقيا يكشف عن العلاقات القائمة بينها.

أما الخطوة الثانية: تحليل الكفايات أو المهارات الرئيسة إلى مكوناتها:

وهناك طرق عدة في تحليل الكفايات من أهمها:

طريقة التحليل الهرمي للكفاية:

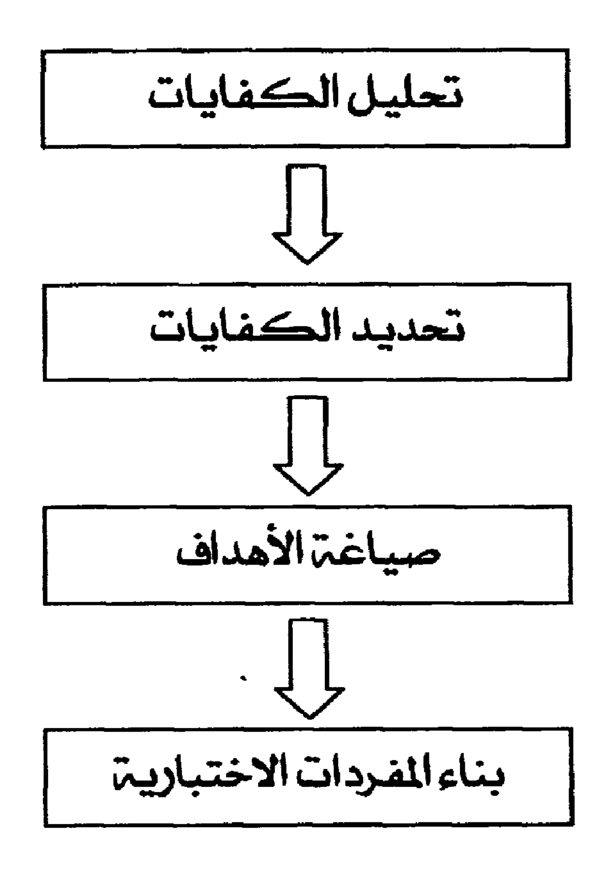
تهتم هذه الطريقة بتحديد المعارف والمهارات المساعدة التي ينبغي تعلمها بترتيب وتتابع بنائي هرمي لكي تتحقق الكفاية أو المهارة الرئيسة المطلوبة.

ولإجراء التحليل البنائي الهرمي لكفاية أو مهارة رئيسة معينة نبدأ بسؤال أنفسنا: ما المتطلبات أو المكونات السلوكية اللازم توافرها لدى المتعلم لكي يحقق الكفاية أو المهارة الرئيسة؟ وبعد تحديد هذه المتطلبات نكرر السؤال نفسه لكل متطلب منها، وبالتالي نحدد متطلبات كل من هذه المكونات السلوكية، وهكذا حتى نصل إلى السلوك المدخلي للمتعلم أي السلوك الذي يكون قد سبق تعلمه، وأكتسبه المتعلم بالفعل. ويتم ترتيب هذه المتطلبات أى المعارف والمهارات المساعدة ترتيبا هرميا بنائيا بحسب أولويات إسهاماتها في تكوين المهارة الرئيسة، حيث يمثل السلوك المدخلى قاعدة الهرم، ثم تتدرج مستويات السلوك من الأبسط إلى الأكثر صعوبة حتى نصل إلى قمة الهرم الذي تمثله الكفاية أو المهارة الرئيسة.

غير إنه من الضروري قبل البدء في إجراء هذا التحليل أن يكون الباحث أو المعلم على دراية بأنماط التعلم الرئيسة ومستوياته المتدرجة ونواتجه لكي يكون التحليل متكاملاً (أنظر الشكل رقم 1).

الخطوة الثالثة: صياغة الأهداف السلوكية:

بعد الانتهاء من عملية تحليل الكفايات أو المهارات الرئيسة للبرنامج التعليمي أو التدريبي، فإن الخطوة الثالثة هي صياغة الأهداف السلوكية المتعلقة بنواتج التحليل. فقد أتضح لنا أن هذه النواتج تتضمن المعارف والمهارات المساعدة التي تسهم في تحققها أو اكتسابها إذا استخدمت طريقة التحليل البنائي الهرمي، بحيث يتم إعادة صياغة هذه النواتج صياغة إجرائية أو سلوكية يمكن ملاحظتها ملاحظة مباشرة وقياسها بمفردات اختباريه. فالاختبار التشخيصي مرجعي الهدف يعتمد في بنائه على الأهداف السلوكية التي تحدد نواتج التحليل التعليمي تحديداً واضحاً، وتعبر عن هذا الأداء تعبيراً دقيقاً لا يحتمل الجدل أو التأويل.



شكل رقم (1) خطوات الاختبارات التحصيلية

لذلك يجب العناية بصياغة الأهداف السلوكية المتعلقة بإجراءات أو مكونات الكفايات أو المهارات الرئيسة التي تم تحليلها. وتشمل عادة عبارات الأهداف السلوكية على أربع عناصر أساسية هي:

1- وصف السلوك المتوقع
2- المحتوى المرجعي
3- شروط الأداء
4- مستوى الأداء

الخطوة الرابعة: بناء المفردات الاختبارية:

وتتطلب هذه الخطوة دراية تامة من جانب الباحث أو المعلم بكيفية انتقاء أنسب أنواع المفردات التي تقيس الأهداف السلوكية المحددة قياساً مباشراً، كما تتطلب التمكن من محتوى البرنامج التعليمي أو التدريبي المعين وفهم خصائص

المتعلمين. وذلك لأن هذه المفردات تستخدم في التمييز بين الذين استطاعوا تحقيق الأهداف المحددة والذين واجهتهم صعوبات، وتشخيص أخطاء وفجوات التعلم.

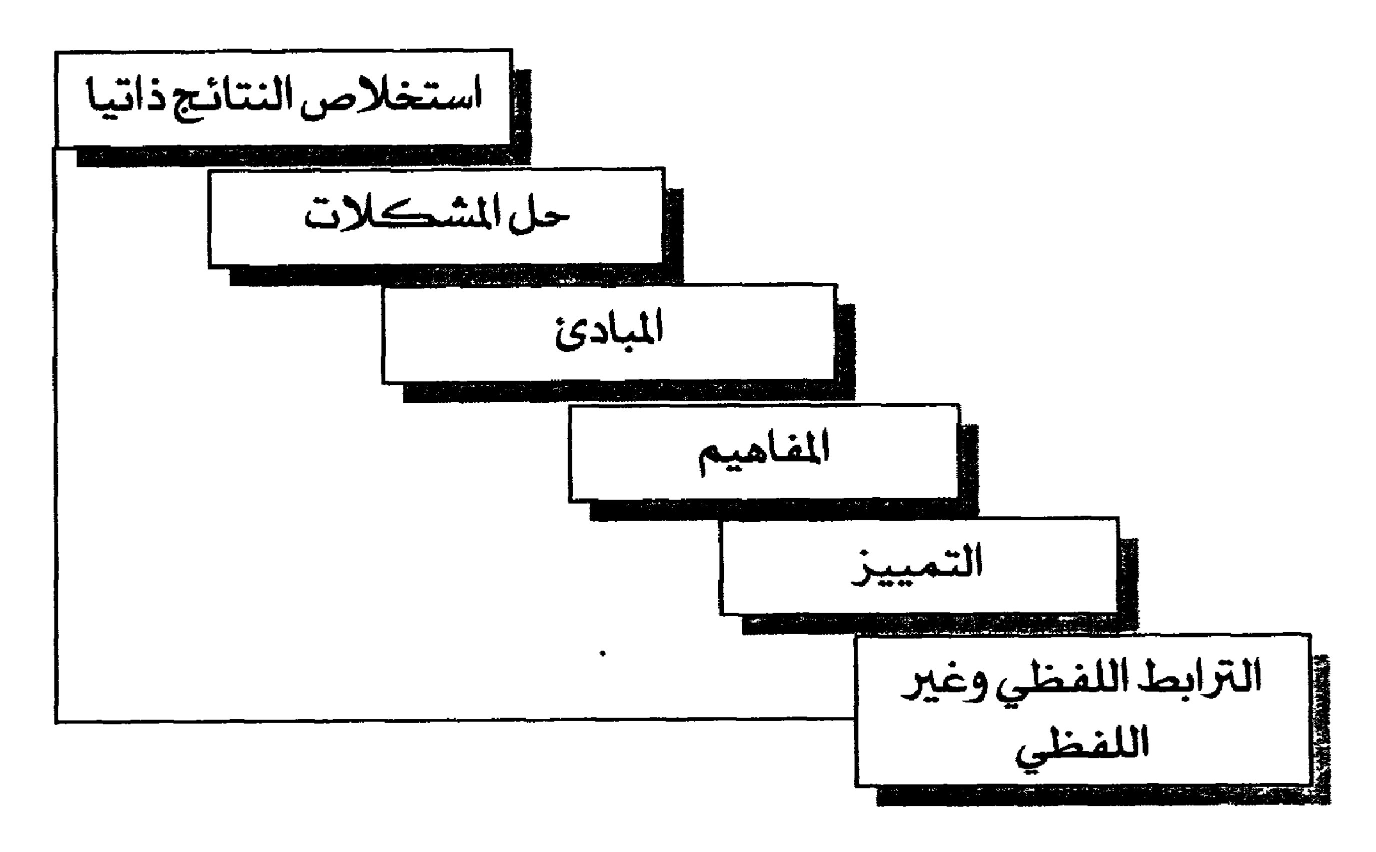
أنماط التعلم ومستوياته:

يصنف جانييه الإمكانيات العقلية في ثلاثة أنماط رئيسة يعتمد كل منها على الآخر وهي: المعلومات، المهارات العقلية، الاستراتيجيات المعرفية، فالمتعلم يتعلم المعلومات ويختزنها في ذاكرته، ويمكن أن تتعلق هذه المعلومات بمادة دراسية أو مجال تعليمي أو تدريبي معين، أو تتعلق بمحتوى يتميز باستمرارية توظيفه في الحياة مثل معرفة الحروف والأعداد والحقائق المتعلقة بالإنسان والبيئة.

ويعتمد المتعلم على هذه المعلومات في اكتساب المهارات العقلية والاستراتيجيات المعرفية. فالمهارات العقلية تتعلق بكيفية أداء نشاط عقلي معين كالتمييز بين مجموعة من الأعداد أو الرموز أو المفاهيم وتطبيق القواعد وحل المشكلات، أما الاستراتيجيات المعرفية فتعد نوعا من المهارات العقلية المتعلقة بسلوك المتعلم بغض النظر عن محتوى مادة التعلم. فهي "إمكانيات" ذات تنظيم داخلي خاص بالمتعلم ويستخدمها دون عون من الآخرين في توجيه عمليات استقبال المثيرات والتذكر والتفكير والإبداع.

وقد حدد جانييه عدة مستويات متدرجة فيما يتعلق بهذه الأنماط أو الإمكانيات الرئيسة الثلاث للتعلم، وكل من هذه المستويات يعتمد على ما يسبقه في إطار البنية الهرمية.

ولكي يتمكن الباحث والمعلم من توظيف هذا التصنيف في إجراء التحليل البنائى الهرمي للكفايات والمهارات الرئيسة سوف نلقى الضوء على كل من هذه المستويات.



شكل (2) مستويات التعلم

1- الترابط اللفظي وغير اللفظي

وهذا النمط الذي يمثل المستوى الأول للتعلم يقصد به العلاقة الترابطية بين مجموعة من المثيرات والاستجابات في تسلسل معين. وقد يكون هذا النمط لفظياً أو غير لفظي. ويتمثل الترابط اللفظي في تعلم الطفل التتابع اللفظي للأعداد حيث يشتمل هذا التتابع على روابط بين المثيرات والاستجابات، كما يتمثل في تعلم اللغات حيث يقوم المتعلم بالربط التسلسلي للكلمات بهدف تكوين جملة أو مجموعة من الجمل، وكذلك سرد المتعلم تواريخ بعض الأحداث في ترتيبها الزمني أو سرد الشعر.

أما الترابط الحركي فيتمثل في المشي الذي يتطلب حركات متسلسلة ومتآزرة بين الرجل والذراع والخصر، أو في فتح باب الغرفة باستخدام المفتاح، أو في تشغيل جهاز معين.

2- التمييز:

وهذا المستوى الثاني للتعلم يعد من المهارات الأساسية وبخاصة لدى الأطفال حيث تكون هناك استجابات متباينة لمثيرات تختلف عن بعضها في صفة أو أكثر، مثل التمييز بين الأشخاص أو أنواع النباتات أو الألوان أو الكلمات الجديدة وغير ذلك.

3- المفاهيم:

ويتعلق هذا المستوى بتعلم المفاهيم واستخدامها، أى تعلم تصنيف المثيرات تبعا لخصائص مجردة كاللون أو الشكل أو الحجم أو العدد. وعندئذ يستطيع المتعلم أن يستجيب بأسلوب واحد لمجموعة من الأشياء التي تكون فئة أو مجموعة معينة. لذلك يعتمد تعلم المفاهيم على تعلم التمييز، وهذا بدوره يعتمد على تعلم الترابط اللفظي وغير اللفظي بين المثيرات والاستجابات.

4- المبادئ:

ويختلف هذا المستوى من التعلم عن المستوى السابق المتعلق بتعلم المفاهيم في إنه يتعلق بتكوين ترابط تسلسلي بين مفهومين أو أكثر في شكل نمط سلوكي استجابة لمجموعة من المثيرات بينما تعلم المفاهيم يتعلق بالتمييز بين سمات أو خصائص مجموعات من الأشياء أو الأحداث أو الأماكن. لذلك فإن تعلم المبادئ التي تربط بين عدد كبير نسبيا من المفاهيم يتطلب تعلم المبادئ الأبسط أي التي تربط بين مفهومين مثلا. ويعتمد هذا أيضاً على مستويات التعلم السابقة.

5- حل المشكلات:

يعد هذا المستوى من مستويات التعلم امتدادا طبيعيا لتعلم المبادئ. فتعلم حل المشكلات يتضمن اختيار مبادئ معينة من مجموعة كبيرة من المبادئ التي سبق تعلمها من أجل حل مشكلة محددة أو تحقيق هدف معين. وفى مواقف حل

المشكلات لا يخبر المتعلم بالمبدأ أو المبادئ التي يستخدمها، بل تقدم له المشكلة وعليه أن يبحث بمفرده عن حل لها. غير إنه يمكن تعلم حل المشكلات من خلال عملية تعليم كيفية تحديد المشكلة وتحليلها، وكيفية اختيار المبادئ المناسبة التي سبق تعلمها وتطبيقها للتوصل إلى الحل الصحيح للمشكلة.

6- استخلاص النتائج ذاتيا:

ويتطلب هذا المستوى من التعلم اكتشاف العلاقات القائمة بين المفاهيم والمبادئ ذاتياً، أي دون تقديم أي عون. ويعد هذا المستوى أعلى مستويات التعلم حيث يصبح المتعلم معتمداً على نفسه ويفكر تفكيراً مستقلاً. وعلى الرغم من أن الإبداع يمكن أن يتضح في جميع مستويات التعلم، إلا أنة يتمثل بدرجة أكبر في مستوى استخلاص النتائج ذاتياً.

الفصل الثامن

الاختبارات التحصيلية

يحتوي الفصل على:

- الاختبارات التحصيلية المدرسية
- الاختبارات التحريرية (المقالية)
- الاختبارات التحصيلية الموضوعية
- الاختبارات المقننة الخارجية
- بناء الاختبارات التحصيلية
- بنك الأسئلة
- اختبارات الأداء

الاختبارات التحصيلية المدرسية:

الاختبارات التحصيلية بشكل عام نوعان: الأول ذلك الذي يعده المعلمون والمدرسون بأنفسهم مثل الامتحانات الشفوية والمقالية، والثاني الاختبارات التحصيلية المقننة والتي تعدها جهة مختصة وقد يكون من المناسب هنا التوسع قليلاً في الاختبارات التحصيلية التي يستعملها المعلمون لتقويم وقياس تحصيل التلاميذ وهي أنماط مختلفة ينبغي أن يجريها المعلم جميعاً دون التركيز على واحدة منها وهذه الاختبارات أو المقاييس هي:

1- الاختبارات الشفوية اليومية:

ربما تكون الاختبارات الشفوية أقدم طريقة استخدمت في تحديد استيعاب المتعلمين للدروس التي تعلموها، فيقال أن سقراط استعمل الاختبارات الشفوية منذ القرن الرابع قبل الميلاد للوقوف على مستوى مستمعيه لكي يبني تعليمه لهم على أساس خبرتهم الماضية.

تهدف هذه الاختيارات إلى التعرف على مدى إتقان التلاميذ للمادة الدراسية بمعزل من قدرات الكتابة والتعبير، وفي تقييم قدرة المتعلم على القراءة والنطق السليم، وتلاوة القران الكريم، وهي ضرورية في المراحل الدراسية الأولى من الدراسة الابتدائية، وكذلك في مجال الحكم على مدى استيعابه للحقائق والمفاهيم، وتستخدم في مناقشة رسائل الدكتوراه والماجستير كما يمكن عن طريق الاختبارات الشفوية الكشف عن أخطاء المتعلمين وتصحيحها في الحال ويستطيع المتعلمين الاستفادة من إجابات زملائهم. وتستخدم الاختبارات الشفوية في اختيار الموظفين، وفحص المظهر والتعبير والحركات والأسلوب الشخصي للأفراد، وتقويم مستوى تأثير المفحوص على الآخرين، وما يتميز به من شخصية من هذه الناحية، تقويم مستوى التنبه الحذر والتبصر لدى المفحوصين. إلا أن جملة من الانتقادات توجه لموضوعيتها وعيوبها الكامنة منها:

أ- تأثر المعلم بالفكرة المسبقة عن الطالب أو تأثر المعلم بحالته المزاجية التي قد تختلف من وقت لأخر.

ب- تأثر استجابة الطالب بالموقف الامتحاني فقد يرتبك التلميذ أمام المدرس ليس لأنه لا يعرف المادة بل لأسباب أخرى منها الخوف من الامتحان نفسه.

ج- حدود الأسئلة الممكن صياغتها لموضوعات محدودة لا تلبث أن تتكرر فيستفيد بعض الطلبة من إجابة الذين تقدموا قبلهم

د- الصدفة واحتمالها فقد يمر على الطالب سؤال أو أكثر تكون سهلة قياساً بأسئلة توجه إلى طالب آخر.

هـ- يتطلب إجراؤها وقتاً أطول مما لا يتيح المجال لإمكانية إخضاع الطلبة كافة إلى الاختبار في الوقت المقرر.

و- يحدث أن يبدأ المعلمون الاختبار باعتماد الدقة والموضوعية ثم لا يلبثون أن يفقدوها بعد اختبار مجموعة من الطلبة بمرور الوقت مما يؤثر سلباً على درجة تقويمهم وثبات تقديراتهم.

ز- عدم شمولية هذه الاختبارات مما لا يفسح المجال لتغطية اكبر قدر ممكن من الموضوعات الدراسية التي يراد قياس تحصيل التلاميذ فيها.

أما أهم مزاياها:

1- تمتاز الاختبارات الشفوية بعدم السماح للغش.

2- توفر فرصة للتعلم من خلال المناقشات التي تدور بين الفاحص والمفحوص.

3- توفر فرصة المواجهة بين الفاحص والمفحوص وجها لوجه وبذلك يتعامل الفاحص مع الإنسان مباشرة بدلا من أن يتعامل معه من خلال الأسئلة المكتوبة على الورقة كما هو الحال في الاختبارات الكتابية المقالية والموضوعية.

4- تناسب الاختبارات الشفوية تقويم بعض الصفات والأهداف التي لا يمكن تقويمها إلا بها.

5- لا تحتاج الاختبارات الشفوية إلى وقت وجهد في إعدادها بعكس الاختبارات المكتوبة.

6- توجد مرونة واسعة في طريقة سبر المعلومات بواسطة الاختبارات الشفوية وذلك بإعطاء الأسئلة السابرة.

وثمة شروط ومواصفات يمكن إتباعها لتحسين الاختبارات الشفوية يمكن تلخيصها بما يأتي:

1- عدم اعتماد الأسئلة الشفوية وحدها في تقدير تحصيل التلاميذ بل ينبغي اعتماد وسائل أخرى منها ملاحظات المعلم على نشاطهم ومجمل الملاحظات المدونة في دفاتر الواجبات البيتية وأحكام المعلمين الآخرين وملاحظاتهم.

2- كتابة الأسئلة الشفوية في أوراق أو قطع مقوى ويدون فيها أكثر من سؤال ويطلب من التلميذ أن يسحب ورقة بصورة عشوائية من بين هذه الأوراق الموجودة داخل الكيس. وغالباً ما تعطى لهم الفرصة في اختيار سؤال من بين الأسئلة المكتوبة.

3- استعمال هذه الاختبارات بشكل محدود يومياً ولمدة قصيرة من الزمن لتحفيز التلاميذ على التحضير اليومي وإكمال واجباتهم أكثر منه لتقويم تحصيلهم.

4- إزالة أسباب الارتباك والتوتر أثناء استجواب التلاميذ خاصة لدى بعض التلاميذ شديدي الحساسية والحرص على إظهار التقبل لديهم كافة.

2- تقويم تحصيل الطلاب من خلال دفاتر تحضيرهم اليومي ونشاطهم في الصف:

ومثل هذا الأسلوب يكمل الأساليب الأخرى خاصة في المراحل الدراسية الأولى ومن مزاياه إنه يعتمد على ملاحظات متعددة تمثل سجلاً عاماً لجهود الطالب ومحاولاته، كما أنها تكشف عن مصادر ضعفه ونوع الصعوبات التي يعاني منها وقد تفيد المعلم في تقويم طريقة تدريسه والموضوعات التي ينبغي التركيز عليها وينبغي أن يأخذ المعلم بنظر الاعتبار مسار تطور الطالب وتحسنه وتجاوزه لصعوباته وأن يقارن الطالب بنفسه لا بالآخرين ويقدر تواصل جهوده ومحاولاته بالنسبة لما كان عليه.

الاختبارات التحريرية (المقالية):

وهي الأكثر شيوعاً في ظروفنا التربوية والأطول تاريخاً وقد تكون أهم وسيلة

تستعمل لتحديد مستوى تحصيل الطالب، ولهذه الطريقة إيجابيات وسلبيات كما لها شروط خاصة لتحسين استعمالها في قياس تحصيل الطلاب.

أ- الخصائص الإيجابية للاختبارات التحريرية (المقالية):

أ - تسمح للطالب بأن يتأمل وينظم أفكاره دون إحراج، أو الإثارة التي قد يسببها موقف الامتحان الشفهي. كما تسمح له بأن يغير ويصحح ويضيف بوقت أكبر مما هو في موقف الامتحان الشفوي.

ب- تنمي القدرة الفكرية للطالب على التعبير وتنظم أفكاره وفهمه للموضوعات.

ج- أقل كلفة في الوقت وتتطلب وقتاً أقل في الأعداد.

د- احتمالات الغش فيها أقل من احتمالاته في الاختبارات الموضوعية.

ب- سلبيات وعيوب الامتحانات التحريرية (المقالية):

أ - ضعف شمولية الأسئلة فقد تأتي الأسئلة من جزء معين من المادة دون شمول الأجزاء الأخرى واحتمال استفادة بعض الطلبة من الصدفة كما قد يتضرر بعضهم الأخر.

ب- ضعف ثبات وموضوعية التصحيح، حيث يتأثر المعلم أو المدرس بعوامل عديدة أثناء التصحيح منها عوامل ذاتية للحالة النفسية للمدرس أثناء التصحيح أو بعوامل تخص أوراق الإجابة كحسن الخط والتنظيم أو تتابع أوراق الإجابات بشكل أو مستوى معين.

ج- عدم وضع تقديرات عادلة ومتوازنة لإجابات الطلبة ويحدث أن يكون الطالب الذي أكثر من الكلام ينال درجة أعلى من طالب آخر عبر عن الإجابة بشكل تام ومختصر.

د- ضعف حساسية المقياس إذ يندر أن يكون مدى الدرجات كاملاً محصوراً بين أعلى درجة ممكنة وأقل درجة ممكنة. فالطالب مهما كتب ينال عليها درجة ما

لذلك يغلب أن تنحصر الدرجات وتتقارب في مدى أقل من المدى الحقيقي للقياس كأن تكون درجات الطلبة محصورة بين 20-90 أو بين 40-85 وهكذا وكلما تقاربت الدرجات فقد المقياس حساسيته للفروق الفردية في التحصيل.

هـ- يختلف معنى الدرجات من طالب إلى آخر فقد ينال الطالب درجة (60) من إجاباته على الأسئلة الصعبة، بينما ينال طالب آخر الدرجة نفسها نتيجة تركيزه على الأسئلة السهلة.

و- ضعف حساسية الاختبار كمقياس نتيجة إضافة أسئلة الترك التي تشجع الطالب على ترك قسم من المادة دون دراستها على أساس إنه لو ظهرت في الامتحان فإنه سيتركها وكلما كثر عدد الأسئلة الترك كلما ضعف الاختبار وتقاربت درجات الطلبة دون دقة، فقد ينال طالب ما 90 أو 100 دون أن يكون قد أستوعب المادة كاملة، بينما ينال طالب أخر الدرجة نفسها وهو متقن لكل مادة دون استثناء.

ز- ضعف صدقها إذ يحتمل أن تقيس الأسئلة معلومات تعتمد على الذاكرة فقط بينما يكون هدف الدرس عمليات عقلية أخرى.

ح- تتطلب وقتاً وجهداً ودقة في التصحيح مما يصعب تحقيقه إذا كان عدد الطلبة كبيراً وهو غالباً ما يحدث في الامتحانات النهائية وحينما تتجمع لدى المدرس أعداداً كبيرة من الدفاتر الامتحانية الأمر الذي يشجع على التسرع والتخمين الذي يضعف من ثبات الدرجات.

ط- ضعف وغموض صياغات الأسئلة فقد تنطوي بعض الأسئلة على غموض في معناه مما قد يفسره الطلبة بأشكال مختلفة وقد يكون المعلم يقصد معنى معيناً في ذهنه لا يكون واضحاً في صياغة السؤال.

ج- شروط ومواصفات الاختبارات التحريرية (المقالية):

1- ينبغي أن تكون الإجابة واضحة ولا تتقبل أكثر من تفسير ويمس الجواب وموضوع واحداً يفهم مباشرة من السؤال.

2- تصمم الأسئلة بشمولية لتحقيق العدالة في توزيع الأسئلة على أجزاء المادة المقصود اختبار الطلبة فيها ودون الاقتصار على جزء معين في أول أو آخر الكتاب أو الفصل.

3- تدرج الأسئلة في الصعوبة إذ ينبغي أن تكون الامتحانات مقياساً متدرجاً من السؤال السهل إلى الصعب إلى الأصعب.

4- عدم وضع أسئلة ترك لأن السؤال الترك يفترض بأن جميع الأسئلة متساوية في الصعوبة وهو يضعف حساسية الاختبار كمقياس للفروق الفردية في التحصيل.

5- تحديد الإجابة النموذجية وتقسيم الدرجات بأوزان مناسبة تبعاً للأهمية النسبية للسؤال أو إلى اعتبارات أخرى يشدد عليها المعلم ويطلبها في الامتحان.

6- توزيع التصحيح وتدقيق جمع الدرجات من قبل أكثر من مصحح إذا توافرت الإمكانيات، كما يحدث في الامتحانات العامة أو أن يعمد إلى أن توزع بحيث يخصص سؤال لكل مدرس ويتولى آخر جمع الدرجات وكلما أزداد عدد المصححين أو أعيد تصحيح الإجابات أزداد ثبات الدرجات وازدادت موضوعية الامتحان.

7- تصاغ الأسئلة بشكل يناسب الوقت المخصص ويشجع الطلبة على الكتابة المنظمة على شكل نقاط أو في ترابط منطقي.

8- تصاغ الأسئلة لتشمل عمليات عقلية وأهدافاً سلوكية مختلفة دون التركيز على أعاده سرد ما حفظ ويفضل أن تكون موزعة باستعمال كلمات عدد، ناقش، قارن، علل، أو وضح الأسباب، أنقد هكذا.

9- يفضل زيادة عدد الأسئلة وتقليص حجم الإجابة عليها بحيث تكون أقرب إلى الأسئلة القصيرة الإجابة وقد تكون الأخيرة أفضل أنواع الامتحانات التحريرية بالنسبة للظروف الدراسية في بعض البلدان النامية، حيث يمكن وضع أكبر عدد من الأسئلة ليكون الامتحان أكثر شمولية للمادة وللأهداف التعليمية، إضافة إلى ضعف احتمالات الغش فيها وصعوبته.

10- تجنب النظر إلى أسماء الطلبة قبل التصحيح وإذا أمكن تغطيتها فتغطى.

11- يفضل تصحيح كل سؤال على حده مرة واحدة لجميع الطلبة قبل الانتقال إلى السؤال الثاني (تجنب تصحيح الأسئلة جميعها للطالب الواحد مرة واحدة).

12- يفضل قراءة عدد معين من الإجابات قبل البدء بالتصحيح ويفضل أخذ عينات تمثل مجاميع للإجابات الممتازة والوسط والضعيفة وقراءة البعض منها قبل التصحيح.

أنواع الأسئلة المقالية

ذكرنا فيما سبق أن الأسئلة المقالية تعطي حرية الإجابة للطالب، ولكن هذه الحرية تتفاوت تفاوتا بينا من سؤال لآخر، فقد يطلب من الطالب أن يعطي جوابا موجزا ومحددا وبدقة في حين قد يعطيه سؤال آخر حرية واسعة في تحديد طبيعة إجابته ومداها. والأسئلة من النوع الأول تسمى "أسئلة ذات إجابات مقيدة" أما الأسئلة من النوع الثاني فهي "أسئلة ذات إجابات موسعة حرة" وفيما يلي توضيح لهذين النوعين:

أ- أسئلة الإجابات المقيدة:

وتضع هذه الأسئلة قيودا على الإجابة المطلوبة، وتقوم بتحديد مضمون المادة المطلوبة بحدودها الضيقة مثل (اجب بما لا يزيد عن سطرين، حدد إجابتك في خمسة اسطر، عدد، عرف، علل، رتب في قائمة).

ويشترط في هذا النوع من الأسئلة ما يلي:

- أن تكون الأسئلة شاملة للمادة جميعها ما أمكن.
- وضع إجابات نموذجية يتم التصحيح على أساسها.
- وضع نقاط رئيسية للأسئلة تحدد مسارات إجابة الطالب
- أن تكون الأسئلة واضحة والتحديد واضحاً.

مزاياها:

- تعطي للطالب الحرية في الإجابة بدرجة اكبر من الأسئلة الموضوعية.
- غالبا ما يطلب في هذه الإجابة الرأي الشخصي، مما يزيد من ثقة الطالب بنفسه.
- تعطي مدى أوسع للإجابة وحريتها.
- تؤدي لإثراء المعرفة نظرا لتنوع الإجابات.

عيوبها:

- أنها أكثر صعوبة في التصحيح من الأسئلة الموضوعية.
- قد يبتعد الطالب في إجاباته عما هو مطلوب، وبالتالي تكون استجابته حول أشياء عامة.
- تتأثر إجابات المفحوص وكذلك التصحيح بذاتية كل من الطالب والمعلم.
- هي ذات قيمة ضئيلة نسبياً لقياس مستوى التركيب والتقويم.
- لا تتيح للطالب إلا فرصة ضئيلة لإظهار قدرته على التنظيم والمكاملة مما يجعلها تقتصر على قياس أهداف من مستويات الفهم والتطبيق والتحليل.

ب- أسئلة الإجابات الحرة

تعطي هذه الأسئلة للطالب حرية غير محدودة في الغالب لتحديد طبيعة إجابته ومداها مما يتيح له الفرصة لإظهار مهارته في التركيب والتقويم مع الضبط الكافي للتأكيد على أن القدرات والمهارات العقلية المطلوبة قد استدعيت من قبل السؤال.

وتمتاز هذه الأسئلة بأنها تشجع وتقيس المكاملة الخلاقة للأفكار والتقويم الكلي للمادة والاقتراب الواضح من حل المشكلة وهذه هي كل أهداف التعلم المهمة التي لا يمكن قياسها بأنواع أخرى من فقرات الاختبارات ولكن هذا النوع من الأسئلة يأخذ وقتا طويلا في التصحيح.

الاختبارات التحصيلية الموضوعية:

ويتضمن هذا النوع من الاختبارات إجابات محددة تتقلص فيها احتمالات الأحكام الذاتية مثل اختيار إحدى الإجابات من بين بدائل عديدة للسؤال الواحد. ولهذا النوع من الاختبارات أنماط متنوعة من الأسئلة ولكل منها استعمالات ومواصفات نسبية، كما أن الاختبار الموضوعي ينبغي أن يجمع بنسبة مناسبة بين أنماط عديدة من الأسئلة الموضوعية وأدناه النسب التقريبية لما يفضل أن تحتوي عليه الاختبارات.

أ- الإجابة القصيرة (محدودة بسطور قليلة) بين 20 - 25٪ من الاختبار الكلي.

ب- الاختيار من البدائل المتعددة ما بين 30 إلى 50٪.

ج- الخطأ والصواب ما بين 10 إلى 15٪.

د- تكملة الجمل أو ملء الفراغات ما بين 10 إلى 15٪.

هـ- المزاوجة بين الكلمات والمصطلحات ما بين 10 إلى 15٪.

و- التعليل وذكر السبب 10 إلى 15٪.

ولهذا النوع من الاختبارات إيجابيات عديدة أهمها شمولها للمنهج المقرر وموضوعيتها (الصدق والثبات) وسهولة التصحيح، كما أن لها سلبيات منها صعوبة التصميم في الوقت والدقة التي تحتاجها ولكل من هذه الأنواع شروط ومواصفات كما لها سلبيات وإيجابيات.

أما بالنسبة للطالب وهي أكثر أثاره له ويتشوق لها، كما أنها لا تستغرق وقتاً طويلاً في الإجابة فالمعروف أن الطالب يستطيع أن يجيب على (100) سؤال من

اختبار الاختيارات المتعددة في مدة ساعة وربع ويمكن أن يجيب على (150) سؤالاً من نوع الصواب والخطأ في ساعة أو أن يجيب على (20) سؤالاً من الإجابات القصيرة في ساعة.

أ- قواعد الاختبارات قصيرة الإجابة:

وتكون هذه الاختبارات في مجالات عديدة جزء من مكونات الاختبار الموضوعي وفي أحيان أخرى تكون مستقلة أي أن تكون أسئلة الاختبار كافة ذات إجابات قصيرة فيقوم المعلم بتقديم عدد كبير من الأسئلة وعادة تغطي هذه الأسئلة المنهج كله ويجيب عليها التلميذ إجابات مقالية قصيرة أو بعبارة قصيرة أو ببعض السطور.

وتتيح هذه الأسئلة فرصة للمعلم لأن يغطي موضوعات عديدة من المنهج أو أهدافاً عديدة من الاختبار، كما تسمح للتلاميذ بأن يجيب بشكل مختصر ومركز على ما أستوعب من مادة، ومجال الصدفة فيها أقل منه في اختبار المقال، حيث يصادف أن يدرس التلميذ جزءاً معيناً من المادة (استناداً إلى تخمينه بما هو مهم) ويجيب على الأسئلة بشكل جيد إذا صادف أن جاءت الأسئلة في ذلك الجزء الذي قرأه، بينما في الاختبار القصير يحتاج الدارس لقراءة المادة كلها.

وبما أن الإجابة قصيرة فإن تصليحها يكون في العادة أسهل وغالباً ما يكون هناك جواب واحد لكل سؤال مما يزيد من دقة توزيع الدرجات وأوزانها الأمر الذي يزيد من ثبات الاختبار قياساً إلى ثبات اختبار المقال. كما يسمح هذا الاختبار بتحليل أخطاء التلاميذ بسهولة.

أما المآخذ التي تؤخذ عليها، ففي مقدمتها الصعوبة النسبية في تصميمها لأنه يتطلب فيه العادة وقتاً أطول وتخطيطاً مسبقاً، كما أن التلميذ قد يركز على حفظه بعض المعلومات والكلمات (المفتاح) دون التعمق الكلي لفهم الموضوع. ففي بعض الحالات يطلب المعلم تحليلاً أو مقارنة يستعمل فيها التلميذ الدروس أو المواد دون سواها، كما أن هذا النوع من الاختبار لا يلائم بعض التلاميذ الذين يكتبون ببطء.

بـ- اختبارات التكميل والاختبارات ذات الإجابات القصيرة

اختبارات التكميل تشبه بدرجة كبيرة الاختبارات ذات الإجابة القصيرة وهما يعتبران وكأنهما نمط واحد من الاختبارات. والفرق الوحيد بينهما أن اختبار التكميل هو عبارة ناقصة ولا يتم المعنى إلا بوضع الكلمة المكملة لها، بينما الاختبار ذات الإجابة القصيرة عبارة كاملة على شكل سؤال ويطلب الإجابة عليها بجواب قصير.

وهذا النمط من الاختبارات هو النمط الوحيد من الاختبارات الموضوعية الذي يتطلب من المفحوص أن يعطي الجواب بدلا من أن يختاره.

مزاياها:

1- تمتاز هذه الأسئلة عن بقية أنماط الأسئلة الموضوعية في أن التخمين فيها اقل من غيرها.
2- تمتاز عن الأسئلة المقالية بأنها أكثر موضوعية وان تصحيحها أسهل.
3- نستطيع أن نغطي بها مساحة اكبر من المادة الدراسية لو قسناها بالاختبارات المقالية التقليدية.

عيوبها:

1- أقل موضوعية من غيرها من أنماط الأسئلة الموضوعية.
2- تتطلب جهدا ووقتا من المصحح في التصحيح نظرا لتعدد الإجابات التي تحتاج لقراءة من قبل المصحح.

قواعد أعداد اختبارات التكميل (ملء الفراغات):

1- من أهم العوامل التي يجب أن تراعى في تصميم الاختبار هي الأهداف التربوية التي تشتمل فيه الاختبار، فبعض الفقرات تختص بتحصيل التلاميذ من المعلومات بينما يكون البعض منها حول تطبيق المعلومات، والبعض لتحليل ومناقشة

المعلومات وبعضها يهتم بالقدرة على تفسير المعلومات والاستنتاج منها وبعضها أيضاً حول قدرات ومهارات أساسية خاصة كالقدرة على القراءة والكتابة.

2- إذا كان القصد من الاختبار قياس تحصيل التلميذ، فالمفروض أن تتعدد الأسئلة بحيث تشمل الأهداف كافة وإذا كان لقياس التحصيل العام للصف فإن الأسئلة تكون اقل عددا.

3- تكتب الفقرة بحيث لا تتطلب إلا إجابة قصيرة ولكن مركزة.

4- يجب أن تكون الفقرة بكلمات ذات علاقة مباشرة بالنقطة الأساسية للهدف السلوكي.

قواعد بناء فقرات اختبارات التكميل (ملء الفراغات):

1- اجعل الفراغ في نهاية الفقرة أو وسطها لا في أولها وهذا يتيح الفرصة للمفحوص أن يقرأ المشكلة قبل أن يصل إلى الفراغ.

2- يجب أن يكون ملء الفراغ بجملة أو كلمة معينة لا يمكن وضع غيرها وآلا يحتمل الفراغ أكثر من أجابه.

3- ابدأ بسؤال مباشر وحوله إلى عبارة ناقصة فاستعمال سؤال مباشر يزيد من احتمال صياغة المشكلة بوضوح حيث إن جواباً واحدا فقط هو المناسب. وكذلك فإن العبارات الناقصة تميل لان تكون اقل غموضا عندما تعتمد على مشكلات صيغت في البداية على شكل سؤال.

4- يستحسن أن لا يكون هناك فراغات كثيرة في كل سؤال (ليس أكثر من ثلاثة فراغات).

5- تجنب أخذ العبارات من الكتب مباشرة لأنه يشجع التلاميذ على الحفظ الآلي.

6- تجنب الفقرات التي تقيس معلومات تفصيلية دقيقة جداً فليس المهم أن يعرف التلميذ أرقاماً تفصيلية بقدر ما يدرك الفكرة الأساسية المجردة.

جـ أسئلة الاختيار من متعدد:

تتكون فقرات الاختيار من متعدد من جزأين: أصل أو مقدمة الفقرة وإبدال متعددة تقدم إجابات ممكنة للمشكلة التي عرضت في الأصل.

وتكون مقدمة الفقرة على شكل سؤال أو على شكل جملة ناقصة يكون الاختيار الصحيح من بين الإبدال متمما لها.

وتشكل الإبدال إجابات ليس فيها إلا إجابة واحدة صحيحة (أو أكثر صحة)، أما الإجابات الباقية فهي إجابات خاطئة (أو اقل صحة) ولكنها مقبولة ظاهريا وتضارب أو تموه على الإجابة الصحيحة المطلوبة، ولذلك يطلق عليها المموهات أو الإبدال والإجابات المموهة، وهي لاشك إنما تنجح في تأدية وظيفتها (التمويه) على الطالب الضعيف أو غير الواثق بنفسه أو غير المتأكد من الجواب الصحيح.

وتتضمن فقرة الاختيار من متعدد – نموذجيا – أربعة أو خمسة إبدال، وقد تقتصر على ثلاثة ابدال، وبالطبع فإن العدد الأكبر من الإبدال يقلل من اثر التخمين، فنظريا في حالة خمسة بدائل توجد فرصة واحدة من خمس لتخمين الجواب الصحيح، بينما في حالة ثلاثة الإبدال تصبح الفرصة واحدة من ثلاث هذا على اعتبار أن الإبدال (الأكبر عددا) اختيرت اختيارا جيدا مع الإشارة إلى صعوبة الحصول على خمسة إبدال كلها فعالة.

وتستخدم فقرات الاختيار من متعدد لقياس أنواع المعارف من مصطلحات وحقائق ومفاهيم وطرائق، كما أنها تصلح لقياس الفهم وعمليات التعلم المركبة فهي قادرة على قياس جميع الأهداف العقلية الإدراكية من معرفة، وفهم، وتطبيق، وتحليل، وتركيب، وتقويم.

مزايا فقرات الاختيار من متعدد:

1- عنصر التخمين فيها اقل من أسئلة الصواب والخطأ.

2- تلفت نظر المفحوص إلى ضرورة التمييز ومعرفة الحقائق معرفة دقيقة.

3- تساعد على قياس قدرة المفحوص في التمييز بين الأحكام الصحيحة والأحكام الخاطئة تمييزا يقوم على الروية والمقارنة وأعمال الفكر.

4- صادقة وثابتة بدرجة اكبر بكثير من بقية الأنواع.

5- توفر فرصة تعلم ممتازة وذلك بسبب من الإجابات المحتملة، وبخاصة حين يكون المفحوص متأكدا من الإجابة الصحيحة.

6- يمكن استخدامها في قياس أهداف معرفية مختلفة كالتذكر والفهم والتطبيق.

7- تعود المفحوص على الحكم الصائب والموازنة وتمييز الأفضل.

8- تساعد في تشخيص أخطاء الطلاب أو سوء الفهم عندهم من خلال استجاباتهم للإبدال الخاطئة.

9- تفرض على الطالب أن يراجع اكبر كمية من المادة المطلوبة بشكل مترو وتوفر له فرصة مناسبة لتفادي تأثير الدراسة السريعة والمكثفة قبل الامتحان.

10- يمكن تحليل نتائجها إحصائيا بسهولة.

11- من السهل تصحيح هذا النوع من الاختبارات وخصوصا إذا استعمل المصحح مفتاح الإجابة المثقب.

عيوب فقرات الاختيار من متعدد:

وعلى الرغم من المزايا المذكورة سابقا عن فقرات الاختيار من متعدد إلا ان عليها بعض الانتقادات وهي:

1- إن إعدادها يحتاج إلى جهد كبير من قبل المعلم ويتطلب منه وقتا طويلا لمراجعة المادة الدراسية لاختيار الفقرات وإعدادها ولا يمكن للمعلم أن ينجح في إعدادها إلا إذا كان قد تدرب كثيرا على كتابة مثل هذه الفقرات.

2- أنها تحتاج إلى نفقات كثيرة في طباعتها ذلك أن كل فقرة منها تتضمن مقدمة وعددا من البدائل مما يتطلب مساحة اكبر وبالتالي تكاليف أكثر للطباعة والتصوير.

3- يحتاج المفحوص إلى وقت كبير في قراءة الأسئلة وفي هذا إهدار لوقته المخصص للامتحان فبدلا من إعطاء كل وقته للتفكير في الإجابة وكتابتها فإنه يعطي جزءا كبيرا من هذا الوقت لقراءة الاختبار الذي يتكون من عدد غير قليل من الصفحات.

4- يظل المجال فيها مفتوحا لشيء من الغش والتخمين.

5- يخشى أن لم تعد إعدادا متقنا أن تكون منخفضة الصدق قليلة الشمول للمادة الدراسية وان تقتصر على قياس هدف التذكر دون غيره من الأهداف المعرفية كما أن الإبدال قد تكون ضعيفة التمويه فيكون البديل الصحيح واضحا.

قواعد أعداد فقرات (الاختيار من متعدد):

1- توزع الفقرات على عدد الوحدات بنسب مناسبة لأهميتها بحيث يغطي الاختبار الأهداف السلوكية وتفاصيل المادة بشكل شامل.

2- لا يمكن تغطية كل المحتوى أو التفاصيل الدقيقة في المنهج أو المادة أنما يكتفي بعينة ممثلة للمحتويات الأساسية منها حسب الأهمية.

3- تصمم كل فقرة لقياس فكرة أو هدف واحد لا يمكن الجمع بين أكثر من فكرة في فقرة واحدة أي تجنب ذكر أداة العطف، (و) للربط بين موضوعين أو هدفين.

4- يجب أن تكون الصياغة واضحة بدون غموض متداخل كأن يحتمل بها أكثر من معنى.

5- يجب أن تكون البدائل الأخرى جذابة وبها شيء يجذب التلميذ الذي لم يستوعب المعلومات بشكل جيد أو حقيقي.

6- يجب أن تكون البدائل مترادفة أو متشابهة من حيث طول الجملة أو شكلها العام كأن تكون جملتين فعلية تبدأ بفعل وجملتين أسمية أو جميعها من نوع واحد.

7- تجنب الإجابة الصحيحة الطويلة بين بدائل خاطئة قصيرة في تركيبها، فالتلميذ قد يختارها بالتخمين لطولها (أي لا تجعلها بشكل يوحي بالإجابة الصحيحة).

8- لا تجعل نوع الإجابة الصحيحة متسقاً في تشابه متقارب كأن يكون في الأكثر في البديلين الآخرين أي جعلها موزعة بشكل عشوائي.

9- أختر أبسط التعابير وبسط المعنى في صياغة البدائل كافة.

10- تجنب أن تكون صياغة البدائل متكررة.

د- أسئلة المزاوجة أو المقابلة

وتسمى أيضاً اختبارات المطابقة وهذا النوع من الفقرات هو صورة معدلة عن نمط فقرات الاختيار من متعدد والفرق بينهما إنه في فقرات الاختيار من متعدد يكون صلب الفقرة عبارة عن مشكلة واحدة ويطلب من المفحوص أن يختار الإجابة من بين قائمة الإبدال المعطاة. أما أسئلة المقابلة فهناك قائمتان: الأولى فيها عدد من المشكلات والثانية فيها إجابات لهذه المشكلات ولكن بترتيب مخالف، والمطلوب من المفحوص أن يربط كل مشكلة من القائمة الأولى، والتي تعرف بقائمة المقدمات، مع جوابها من القائمة الثانية، والتي تعرف بقائمة الإجابات أو الاستجابات، ويكون ذلك بأن يكتب أمام كل مشكلة من القائمة الأولى رقم الإجابة التي وردت في القائمة الثانية.

د- استخدامات فقرات المزاوجة:

تستخدم أسئلة المزاوجة لقياس حقائق ومعلومات مترابطة في وقت قصير نسبياً ولكن استعمالها ينحصر في موضوعات يمكن أن تشتمل على بنود متجانسة تقترن بخصائص معينة وهذه تؤلف في الغالب معلومات يسهل تذكرها.

لكن من الممكن أن تقيس أسئلة المزاوجة عمليات مركبة عندما يكتب عدد من أسئلة الاختيار من متعدد ويكتشف أنها تتناول بنودا متجانسة فتحول إلى أسئلة مزاوجة.

وتفيد أسئلة المزاوجة بشكل كبير في اختبار معاني المفردات وتواريخ الأحداث ونسبة الكتب إلى مؤلفيها والأحداث إلى عواملها والنظريات إلى أصحابها والاكتشافات إلى مكتشفيها والقادة إلى المعارك التي قادوها والدول إلى عواصمها والرموز وما تدل عليه كالرموز الكيمياوية وأسماء العناصر التي تدل عليها وهكذا.

مزاياها:

1- سهولة إعدادها والاختصار في النفقات في الورق المستعمل إذا ما قورنت بفقرات الاختيار من متعدد.

2- انخفاض فرصة لجوء المفحوص إلى التخمين بالنسبة لغيرها من الأسئلة الموضوعة الأخرى.

3- توفير الجهد على المعلم نتيجة لاستخدام قائمة من المشكلات معها قائمة واحدة من الاستجابات في حين كان يحتاج لإعداد قوائم من الإبدال أو الإجابات لكل مشكلة.

4- توفير الجهد على المفحوص فبدلا من أن يقرا عددا من الإبدال للإجابة عن سؤال واحد فإنه يقرا في حالة المزاوجة عددا مقاربا من الإبدال ليجيب عن عدد من الأسئلة لا عن سؤال واحد فقط.

5- تستعمل في الربط بين الأشياء المتماثلة العلاقة.

6- يمكن أن نستبدل بقائمة الاستجابات اللفظية موضوعات أخرى كالصور أو الخرائط أو الرسوم البيانية.

عيوبها

1- لا تقيس هذه الأسئلة القدرة على الفهم والتمييز والمحاكمة عند المفحوص.
2- يلاقي واضع الأسئلة صعوبة في بعض الأحيان في إيجاد الأعداد اللازمة من المواد المترابطة والمتجانسة اللازمة لمثل هذه الأسئلة.
3- لا تقيس إلا جوانب محدودة تتمثل في العلاقة بين عنصر وآخر.

قواعد فقرات المزاوجة:

1- يجب أن تكون أزاوج الفقرات متجانسة ومتعلقة الواحدة بالأخرى بشكل واضح.
2- أختصر من طول الفقرة وتفضل المصطلحات والأسماء والعناوين.
3- يفضل تنظيمها بشكل عمودي متقابل.
4- تكون أعداد المصطلحات والمفاهيم غير متكافئ أي أن تكون أحد الأعمدة الأكثر أو أقل من العمود المقابل بحيث يبقى البعض منها بدون مزاوجة لزيادة ثبات الإجابة وتقليل أثر التخمين.

هـ – أسئلة الصواب والخطأ

يتألف اختبار الصواب والخطأ من عدد من العبارات بعضها صحيح وبعضها خاطئ ويطلب من المفحوص أن يذكر إذا كانت العبارة صواباً أم خطأً، فإذا كانت العبارة صواباً فإنه يذكر إلى جانبها كلمة (نعم) أو (صواب) أو (صح) أو (العبارة صحيحة) أو قد يضع إشارة (✓) أو يضع دائرة أو إشارة (✗) على الحرف ص إذا كانت العبارة صحيحة وذلك حسب ما يطلب في السؤال. أما إذا رأى المفحوص أن العبارة خاطئة فإنه يذكر إلى جانبها كلمة (لا) أو (خطأ) أو (العبارة خاطئة) أو قد يشير بعلامة (✗) أمام العبارة، أو يضع دائرة أو إشارة (✗) على الحرف خ وذلك حسب ما يطلب منه في السؤال.

وتبدو هذه الاختبارات لأول وهلة سهلة التحضير، ولعل ذلك راجع إلى أن كثيرا من المعلمين يلجأون إلى اخذ عبارات من الكتاب المقرر كما هي بدون تعديل، ويحورون بعضها الآخر تحويرا بسيطا لكي تبدو خاطئة، ويؤلفون منها أسئلة الصواب والخطأ. فبعض هذه الأسئلة يكون من الوضوح بحيث يصيب في الإجابة عليها معظم المفحوصين، وبعضها الآخر يكون على درجة من الغموض بسبب التحوير المصطنع الذي ادخل عليها، بحيث يلتبس معناها حتى على المتفوق منهم. ومن الجدير بالذكر أن كتابة سؤال ركيك من أسئلة الصواب والخطأ هو أمر سهل فعلا، لكن كتابة سؤال جيد لا غموض فيه ولا التباس يتطلب درجة عالية من المهارة.

مجالات استعمال أسئلة الصواب والخطأ:

يستخدم هذا النوع من الاختبارات لقياس أهداف تكاد تنحصر في المعلومات التقريرية وفي قياس أهداف معرفية بسيطة من فئة التذكر والحفظ، ولكن إذا أعدت أعدادا جيدا واعتني بهذا الإعداد فإنها يمكن استخدامها في قياس قدرة التلميذ على التمييز بين حقيقة ورأي أو على التعرف على صحة علاقة سببية.

أنواع أسئلة الصواب والخطأ:

تكتب أسئلة الصواب والخطأ بأشكال مختلفة أكثرها انتشارا هو أن توضع عبارات ويطلب من كل مفحوص أن يضع إشارة (✓) إذا كانت العبارة صحيحة وإشارة (✗) على العبارة الخاطئة، أو أي كلمة أو إشارة تدل على صحة العبارة أو خطئها.

وهناك نوع من أسئلة الصواب والخطأ يوضع أمام كل فقرة أو عبارة ثلاثة اختيارات هي: صحيحة، خاطئة، لا اعرف، وفي هذا النوع يضع المفحوص إشارة على الاختيار الذي يدل على صحة العبارة إذا كان يعتقد أنها صحيحة، أو إشارة على الاختيار الذي يدل على خطأ العبارة إذا اعتقد أنها عبارة خاطئة. أما إذا لم

يكن متأكدا من صحتها أو خطئها فإنه يشير إلى الاختيار الذي يبين إنه لا يعرف الجواب وذلك لكي يحمي نفسه من عقوبة تخمين.

أما النوع الثالث من أسئلة الصواب والخطأ فيطلب فيه من المفحوص أن يبين ما إذا كانت العبارة صائبة أم خاطئة، وان يصحح ما فيها من خطأ إذا كانت خاطئة وقد يطلب أن يذكر السبب الذي جعل العبارة خاطئة.

وهناك نوع من أسئلة الصواب والخطأ يكون فيه قائمة متكاملة من العبارات ويطلب من المفحوص أن يبين ما فيها من عبارات صحيحة، ويسمى هذا النوع من الأسئلة بالأسئلة العنقودية.

مزاياها:

ميزة هذه الأسئلة هي شمولها لجزء كبير من المقرر الدراسي في مادة الاختبار.

عيوبها:

1- التخمين: تلعب الصدفة والتخمين دورا كبيرا في هذا النوع من الأسئلة، إذ أن احتمال أن يصيب المفحوص في الإجابة بمحض الصدفة ودون أن يستند إلى معرفة يقينية هو بنسبة 50٪ وقد ترتفع نسبة التخمين عن ذلك إذا اقتطعت من الكتاب المقرر عبارات وأضيفت إليها مموهات مصطنعة لتجعلها تبدو صائبة أو خاطئة هذا ويمكن تقليل التخمين بزيادة فقرات الاختبار.

2- يشجع هذا النوع من الأسئلة على التعلم من غير فهم أكثر مما يشجع على التفكير النقدي، وعليه، فإن الاختبار الذي يحوي أسئلة من هذا النوع ضعيف من حيث نوع التعلم الذي يشجع عليه ومن حيث النتيجة التربوية التي ينتهي إليها.

3- هذا النوع من الأسئلة قد يعرض المفحوص لتعلم معلومات غير صحيحة كونه يضم عبارات خاطئة قد ترسخ في ذاكرة المفحوص.

4- تكون الأسئلة غامضة ومضللة.

5- قدرة هذه الأسئلة على التشخيص ضعيفة، كما ان استجابة المفحوص عليها لا تعبر عن حقيقة أدائه، فهي منخفضة الصدق.

6- عامل الثبات لهذا النمط من الأسئلة منخفض.

قواعد فقرات الخطأ والصواب:

1- يجب أن تكون كل فقرة تمثل فكرة معينة واحدة ولا يجوز الجمع بين أكثر من فكرة أو مصطلح أي تجنب استعمال (و).

2- يجب أن تصاغ الفكرة بشكل قاطع بحيث لا تتحمل إلا أجابه واحدة صحيحة أو خاطئة.

3- يجب أن تكون العبارات موجبة لا منفية، وذلك لان كثيرا من المفحوصين لا يلتفتون للنفي، وقد يقرءون العبارة المنفية خطأ، وإذا كان لا مفر من استخدام عبارة منفية فيحسن وضع خط تحت أداة النفي للفت نظر المفحوص إليها. كما يمكن تجنب شجع التلميذ على الإجابة عن الأسئلة السهلة أولاً وقبل غيرها، ومن ثم العودة إلى الفقرات الصعبة. صياغة فقرات ذات النفي المتداخل أو المزدوج.

4- تجنب كلمات التعميم مثل غالباً، دائماً، أحياناً، نادراً، كثيرا، في معظم الأحيان.

5- استخدام صياغات لغوية قصيرة وواضحة وغير معقدة.

6- تجنب فقرات التي توحي بالإجابة الصحيحة.

7- أكثر من عدد الفقرات وكلما أزداد عدد الفقرات ضعف احتمال الصدفة.

8- أجعل الفقرات مستقلة بعضها عن بعض بحيث لا تجيب البعض منها عن غيرها في مكان آخر.

9- يفضل تحديد وقت مناسب تقديري للإجابة عن الأسئلة كافة دون أن يجبر التلميذ على الالتزام به، وذلك لمساعدته على توزيع وقت الإجابة.

10- يفضل تعديل العلامة التي يحصل عليها الطالب في هذا النوع من الأسئلة عن طريق معالجة اثر التخمين التي تتم بحسب المعادلة الآتية:
العلامة النهائية المعدلة = عدد الإجابات الصحيحة – عدد الإجابات الخاطئة

الاختبارات المقننة الخارجية

هذه الاختبارات تصمم وتبنى من قبل متخصصين في هذا المجال وتستعمل ضمن ضوابط يحددها دليل خاص بكل اختبار، كما يشرف على تطبيقها أشخاص مدربون لهذا الغرض، وتعد هذه الاختبارات كأداة يمكن استعمالها لأغراض مهمة منها قياس تحصيل التلاميذ والتعرف على مستوياتهم، كما يمكن استعمالها لأغراض أخرى يحددها الدليل الخاص بها وفي الغالب فإن تصميم الاختبارات تبعاً للاستعمالات المطلوبة.

وبالإمكان عمل اختبارات تحصيلية لكل مادة على حده أو باختصار جماعي لأكثر من مادة أي تصميم اختبار واحد للمعلومات العامة يتضمن فقرات المواد الدراسية المتقاربة في وظائفها أو أنماطها مثل درس الصحة العامة، والجغرافية والتاريخ والزراعة ... الخ.

ولما كانت هذه الاختبارات تصمم من قبل متخصصين فأننا نختصر أدناه أهم شروط ومواصفات الاختبار المقنن.

شروط ومواصفات الاختبار المقنن الجيد:

- أن يكون الاختبار صادقاً في قياس ما صمم لأجله.
- أن يكون الاختبار ثابتاً لا تتغير درجاته في مرات الإجراء المختلفة.
- أن تكون فقراته صادقة بحيث يتفق المختصون على صدق ما تقيسه.
- أن يكون حساساً للفروق الفردية بحيث يشتت الدرجات بما يقارب منحنى التوزيع الطبيعي ويقيس أغلب المستويات.

- أن تكون تعليماته وشروط تطبيقه مقننة أو موحدة للجميع.
- أن تكون له معايير مشتقة من عينات ممثلة للمجتمع الأصلي.
- أن يقيس عاملاً أو صفة أو ظاهرة مستقلة نسبياً.
- أن يكون عملياً سهل التطبيق، اقتصادياً بسيط الكلفة وملائما في الوقت الذي يستغرقه.
- أن تكون نسبة ثباته وصدقه قد تجاوزت حد الخطأ المعياري.
- أن يحدد له خطأ محتمل وخطأ معياري.
- أن تكون فقراته ممثلة للأهداف المراد قياسها أو مستمدة من تحليل مضمون موضوع أو مادة أو من تحليل طبيعة أداء أو عمل أو مهنة.

قواعد بناء الاختبارات التحصيلية:

هناك مجموعة من الخطوات الأساسية التي يجب إتباعها عند بناء الاختبارات التحصيلية لتكون عاملا مساعداً في قياس النواتج التعليمية، وتحقق أكبر قدر من الأهداف التربوية المحددة، وهذه الخطوات أصبحت محل اتفاق بين معظم المختصين في مجال القياس والتقويم وهي كالتالي:

الخطوة الأولى: تحديد الغرض من الاختبار

والمقصود أن يقوم المعلم في هذه الخطوة بتحديد غرض (وظيفة) الاختبار الذي سيطبقه - ومع أن الغرض العام هو قياس التحصيل الدراسي - إلا ان هناك أغراضاً أخرى للاختبارات، فمنها التشخيصي الذي يستخدم للتعرف على الصعوبات التعليمية التي يعاني منها الطلاب في مجال معين وتحديد أسباب ذلك، أو التكويني (البنائي) لتقديم تغذية راجعة للطلاب وللمعلم عن مدى التقدم في عملية التعلم، أو الاختبار الختامي (النهائي) لإعطاء درجات أو شهادات تبين مدى تحقيق المستوى المطلوب من التحصيل في نهاية المدة المخصصة لتدريس المادة.

إن جميع تلك الاختبارات تخدم الغرض العام، ولكونها الوسيلة المستخدمة لخدمة أغراض التقويم بأنواعه المختلفة فقد أعطيت هذه الاختبارات نفس مسميات أنواع التقويم.

الخطوة الثانية: تحديد أهداف الاختبار

وهذه الخطوة ترتبط بالنواتج التعليمية التي يقيسها الاختبار التحصيلي، والتي تم تحديدها بواسطة أهداف سلوكية تعكس بدقة المطلوب تحقيقه من خلال هذا الاختبار، وخطوة تحديد الأهداف تتطلب شرطين:

1- التعرف على أنواع الأهداف التعليمية ومستوياتها.
2- صياغة الأهداف التعليمية للمادة بعبارات سلوكية تعكس نواتج تعليمية يمكن ملاحظتها وقياسها.

وهذا يأتي تأكيدا على أن العملية التعليمية – العلمية ينبغي أن يخطط لها وتنفذ وتقوم في ضوء أهداف تعليمية محددة.

الخطوة الثالثة: تحليل المحتوى

ويقصد بذلك تحليل محتوى المادة الدراسية وتحديد مواضيعها ومفردات كل موضوع سيغطيه الاختبار، بحيث يكون هذا التحليل معتمداً على تقسيمات مناسبة وذلك من أجل ضمان تمثيل فقرات الاختبار لمفردات محتوى المادة تمثيلاً متوازناً في تحقيق الشمول للاختبار (صدق المحتوى) والذي يعتبر أهم صفة من صفات الاختبار الجيد.

الخطوة الرابعة: إعداد جدول المواصفات

جدول المواصفات يصف ويحدد الموازنة بين أنواع السلوك المراد تحقيقه (الأهداف) والمحتوى يعتمد في ذلك على أهمية كل موضوع في الكتاب المقرر.

كيف يتم بناء جدول المواصفات:

جدول الموصفات هو مخطط تفصيلي يربط العناصر الأساسية للمحتوى بمجالات التقييم ومهاراته الفرعية، ويحدد الأهمية النسبية لكل منها. أي بمعنى إنه يقيس مدى تحقق صدق المحتوى، والصدق يحتوي على عنصرين أساسيين هما: الشمولية والتمثيل. والشمولية تعني يجب أن تكون فقرات الاختبار تشتمل (تغطي) على جميع مكونات المحتوى من أهداف. أما التمثيل فيعني أن تكون فقرات الاختبار تمثل عينة المحتوى تمثيلا صادقاً لجميع جوانب التحصيل. لذلك من الضروري جداً تحليل المحتوى قبل كل شيء.

وبذلك يمكن القول أن جدول المواصفات يحقق عدة فوائد منها:

1- يعطي صدقا كبيرا للاختبار.

2- يعطي المتعلم الثقة بعدالة الامتحان.

3- يعطي كل جزء من المادة الدراسية الوزن الحقيقي لها

4- يساعد في قياس مدى تحقق أهداف المادة بدرجة اكبر

ولبناء جدول المواصفات نتبع الخطوات الآتية:

1- تحديد الأهداف التعليمية للمادة الدراسية التي يسعى المعلم لمعرفة مدى تحققها وتختلف المواد الدراسية من حيث تركيزها على نوعية معينة من الأهداف.

2- تحديد العناصر التي يراد قياسها في المادة الدراسية.

3- تحديد نسبة التركيز لكل جزء في المادة الدراسية وذلك من خلال معرفة عدد الحصص المقررة للوحدة الدراسية مقسومة على عدد الحصص الكلية الدراسية مضروبة في 100٪ أي أن:

$$\text{نسبة التركيز} = \frac{\text{عدد حصص الوحدة الدراسية}}{\text{عدد حصص جميع المادة الدراسية}} \times 100$$

4- تحديد نسبة الأهداف من المستويات المختلفة ويتم هذا من خلال التركيز على هذه الأهداف في أثناء عملية التدريس.

5- تحديد عدد أسئلة الاختبار المراد وضعها للامتحان.

6- تحديد عدد الأسئلة لكل جزء من المادة وذلك حسب المعادلة التالية:

عدد الأسئلة لكل جزء = عدد الأسئلة الكلي × نسبة التركيز × نسبة الهدف

طرق تحديد أوزان الوحدات الدراسية في جدول المواصفات

الطريقة الأولى

جد عدد الأهداف في الوحدة الدراسية، ثم أقسمه على عدد الأهداف الكلية للكتاب (الفصل الدراسي) واضرب النتائج في 100٪

مثال:

رقم الوحدة في الكتاب	1	2	3	4	المجموع
عدد أهداف الوحدة	10	25	15	20	70
وزن الوحدة	٪17	٪42	٪25	٪33	٪100

الطريقة الثانية:

جد عدد صفحات الوحدة الدراسية ثم أقسمه على مجموع صفحات الكتاب (الفصل الدراسي) وأضرب الناتج في 100٪

مثال:

رقم الوحدة في الكتاب	1	2	3	4	المجموع
عدد صفحات الوحدة	35	50	40	30	155
وزن الوحدة	٪23	٪32	٪26	٪19	٪100

الطريقة الثالثة:

جد عدد الحصص المقررة لتدريس الوحدة الدراسية ثم أقسمه على مجموع الحصص المقررة (الفصل الدراسي) واضرب الناتج في 100٪

مثال:

رقم الوحدة في الكتاب	1	2	3	4	المجموع
عدد الحصص	8	15	12	10	45
وزن الوحدة	18٪	33٪	27٪	22٪	100٪

ويمكن للمعلم إيجاد وزن وحدة دراسية تبعاً لأهمية هذه الوحدة ومراعاة ذلك عند حساب أوزان الوحدات المختلفة بالطرق السابقة وزيادة نسبة مئوية بسيطة لوز الوحدة الأكثر أهمية.

فوائد جدول المواصفات

- يوزع فقرات الاختبار لتشمل أنواعا مختلفة من المجالات ومهاراتها الفرعية.
- يوزع فقرات الاختبار لتشمل الموضوعات كافة.
- يوفر صدقاً عالياً للاختبار.
- يجعل الاختبار أداة تشخيصية علاوة على كونه أداة تحصيلية.
- يوزع الزمن على الموضوعات وأهميتها فيعطي الوزن الحقيقي لكل جزء من أجزاء المادة.

جدول المواصفات للاختبار

الرقم	اسم الوحدة	الوزن 100٪	الدرجة	مجالات التقويم			المجموع 100٪
				المعرفة والفهم ٪	٪	٪	

الخطوة الخامسة: كتابة أسئلة الاختبار

وتتضمن هذه الخطوة تحديد نوع فقرات الاختبار وذلك بعد تحديد عددها في الخطوة السابقة، بحيث تناسب قياس الأهداف التي حددت مسبقاً وذلك بناء" على عدد من العوامل منها: الأهداف المراد قياسها، طبيعة المحتوى، مهارة المعلم في وضع الأسئلة الوقت المخصص للاختبار، عدد الطلاب في الصف، المكان الذي سيؤدي فيه الاختبار، الإمكانيات المتوفرة للتصحيح ... الخ.

جميع العوامل السابقة ينظر إليها بصورة كلية، ثم يتخذ القرار باختيار نوع معين من الفقرات (الأسئلة) المناسبة للاختبار، ويستحسن كتابة أسئلة تزيد عن الحاجة الفعلية، حيث يتوقع أن لا تكون جميع هذه الأسئلة صالحة أو مستوفية للشروط، وبذلك يتوفر للمعلم فرصة للاختيار والحذف من بين تلك الأسئلة في الخطوة اللاحقة.

الخطوة السادسة: مراجعة أسئلة الاختبار

في هذه الخطوة تتم مراجعة الأسئلة بعد تجميعها وكتابتها ووضعها في صورة اختبار، بحيث يتم تعديل أو حذف أو إضافة بعض الفقرات عند الحاجة لذلك،

لتكون مهيأة للاستخدام بشكل يضمن وضوحها لجميع الطلاب من جهة، وتمثيلها للنواتج التعليمية المطلوبة من جهة اخرى.

الخطوة السابعة: ترتيب أسئلة الاختبار

والمقصود هنا أن ترتب الأسئلة بشكل يسهل على الطلاب قراءتها وتتبعها وفهمها، ومن ثم الإجابة عليها بطريقة منظمة ومرتبة تيسر على المعلم تصحيحها وتقدير درجاتها بدقة وموضوعية، كما أن ترتيب الأسئلة كذلك سيقلل من كثرة استفسارات الطلاب حولها ويجعلهم يستثمرون الوقت المخصص للاختبار في الإجابة على أسئلته فقط.

وهناك عدة طرق لترتيب فقرات الاختبار حيث أنها ترتب حسب نوع الفقرة، أو الصعوبة، أو المحتوى، وكل ذلك يمكن تحقيقه بطريقة علمية ميسرة، بمعنى أن تكون الفقرات المتماثلة مع بعضها، ومن حيث الصعوبة فإنها ترتب من الأسهل إلى الأصعب. أما المقصود بالترتيب حسب المحتوى فإن ذلك يعني تسلسل الفقرات في ورقة الاختبار بشكل يتوافق مع تسلسل محتوى المادة الدراسية، ويمكن ترتيب الفقرات حسب نوعها وصعوبتها وحسب تسلسل المحتوى في وقت واحد.

الخطوة الثامنة: إخراج الاختبار بصورته النهائية

ويقصد بذلك إخراج أوراق أو كراسة الاختبار بصورة جيدة من حيث الطباعة والمراجعة ووضع التعليمات اللازمة والمحددة لكيفية الإجابة المطلوبة، وكل ما يتعلق بضمان معرفة جميع الطلاب بمختلف مستوياتهم لما هو مطلوب منهم على وجه التحديد، لكي لا يحدث أي إشكال يتعلق بورقة الأسئلة أثناء فترة الاختبار، وكذلك توزيع الدرجات بشكل مفصل تسهل معه عملية تقدير الدرجات (التصحيح) بأكبر قدر ممكن من الدقة والموضوعية وبعيد عن الذاتية.

بنك الأسئلة

المفهوم

بنك الأسئلة هو مكان آمن توضع فيه مجموعات متنوعة من الأسئلة ذات مستويات مختلفة في تقدير قدرة معينة ويسهل عن طريقه سحب أو إضافة مجموعة أو عدد معين من الأسئلة المختلفة المقننة أي التي لها خصائص سيكومترية مميزة ومعلومة مثل معامل السهولة والصعوبة ومعامل التمييز وفعالية المشتتات وكذلك صدق وثبات الفقرات والمصنفة على وفق وحدات المقرر الدراسي وحسب المستويات العقلية المعرفية المطلوب أداؤها أثناء الإجابة عليها، وذلك بطريقة تشبه إلى حد ما تنظيم وفهرسة الكتب. وهذا التصنيف يمكننا من معرفة كل سؤال. وتخزن الأسئلة في ذاكرة الحاسب الآلي على وفق برنامج سابق التجهيز، ويتم معايرتها عن طريق نماذج معينة وباستخدام برامج خاصة بها.

الأهداف العامة والخاصة من أنشاء بنك الأسئلة:

هناك هدف عام لإنشاء بنك الأسئلة يتمثل في تطوير أدوات تقويم التحصيل الدراسي باعتبار التقويم أحد العمليات الهامة لتطوير المنظومة التعليمية كما توجد أهداف خاصة لبنك الأسئلة تتلخص في الآتي:

1- تحسين نوعية الأسئلة من حيث الشكل، وضمان جودة الخصائص السيكومترية لها.

2- ضمان استخدام الأسئلة الجيدة أكثر من مرة، وبالتالي ضمان مستوى جيد من الاختبارات بصفة مستمرة من سنة إلى أخرى والمحافظة على مستوى الاختبارات كل عام.

3- إعداد كوادر من واضعي الأسئلة.

4- تزويد المعلمين بمجموعة من الأسئلة المقننة تحمل أفكاراً جديدة يمكن أن يستعينوا

بها أثناء عملية التدريس، أو أثناء عملية التقويم البنائي المستمر لكل جزء من أجزاء المحتوى.

5- تدريب المعلمين على وضع أسئلة جديدة مشابهة لأسئلة بنك الأسئلة.

6- توفير وقت وجهد المعلم الذي يبذله في بناء أسئلة الاختبارات التحصيلية الموضوعية واستغلال هذا الوقت في تحسين عملية التدريس.

7- مساعدة الطالب على التعلم الذاتي والاهتمام بالتعرف على نتائج التعلم لكل طالب على حدة وفق سرعته وإمكاناته وليس التعرف على الدرجة الكلية لنتائج الاختبار.

8- التخلص من مشكلة سرية الاختبارات سواء عند وضع الأسئلة أو عند التداول أو عند التطبيق.

9- استخدام بنك الأسئلة يساعد المعلم على مقارنة نتائج أداء الطلاب في سنة (ما) بنتائج زملائهم الذين سبقوهم، ويساعد على تحديد أسباب الارتفاع في معدلات الرسوب إذا كان راجعا إلى مستوى أداء الطلاب أو إلى مستوى صعوبة الأسئلة.

اختبارات الأداء

اختبارات الأداء هي تلك الاختبارات التي تقيس أداء الطالب بهدف تعرف بعض الجوانب الفنية في المنهاج وتعرف بعض المهارات التي لا يمكن قياسها بالاختبارات الشفوية أو الكتابية وبذلك فهي لا تعتمد على الأداء اللغوي المعرفي للطالب وإنما تعتمد على ما يقدمه الطالب من أداء عملي في الواقع ولتوضيح ذلك نقول إنه يمكن أن يدرس الطلاب مواد نظرية تعليمية مختلفة مثل مواد علم النفس التربوي وعلم نفس النمو والقياس والتقويم وغير ذلك. وقد يجتازون هذه المواد نظريا بنجاح ولكن الفهم الحقيقي لهذه المواد إنما يكون في القدرة على تطبيق مفاهيم هذه المواد في موقف عملي كأن يكون إعطاء الطالب حصة صفية بنجاح.

مجالات استخدام اختبارات الأداء

تستخدم الاختبارات الأدائية في معظم مجالات المواد الدراسية، لأهميتها في قياس مدى إتقان هذه المواد وفهمها. وبشكل عام يمكن حصر بعض هذه المجالات فيما يلي:

1- التحقق من إتقان الطالب للمهارات المرتبطة بالعلوم الطبيعية كالكيمياء والفيزياء والأحياء، فمن خلال تعرف الطالب المركبات الكيميائية وإجراء التجارب المختلفة في الكيمياء والفيزياء والأحياء يمكن الاستدلال على إتقانه لهذه المهارات.

2- برامج التدريب المهني، كتعلم الضرب على الآلة الكاتبة، وأعمال السكرتارية، والعزف على الآلات الموسيقية، فليس هناك طريقة أفضل لمعرفة قدرة الطالب على الكتابة على الآلة الكاتبة، من مراقبة ذلك عملياً وكذلك العزف وغير ذلك.

3- المدارس الفنية والمهنية كمدارس التمريض والمدارس الصناعية والتجارية حيث تعتبر الاختبارات الأدائية جزءا لا يتجزأ من الامتحان النهائي. وتعتمد هذه المدارس أساسا على التدريب العملي لاستكمال متطلبات التخرج.

4- تعلم اللغة، وتعتمد مدارس تعليم اللغات، والمدارس الثانوية، والمعاهد، والجامعات، التدريب العملي، كمتطلب للتخرج في مجال اللغات. ومن اجل ذلك أنشيء في كثير من هذه المراكز العلمية ما يسمى بمختبر اللغة، وذلك من اجل التدريب على صوتيات اللغة وبالتالي يستخدم الاختبار الأدائي للتحقق من إتقان الطالب لما تدرب عليه.

5- المراكز التعليمية التي تقدم خبرات مباشرة إلى المجتمع، ونخص بالذكر كليات الطب والهندسة والتمريض، ومعاهد وكليات المعلمين، فالطالب في كلية الطب يأخذ مالا يقل عن ربع مواده عمليا وبالتالي يتم اختباره أدائيا في نهاية هذه

المواد. وكذلك بالنسبة لطالب الهندسة أو التمريض. وبخصوص إعداد المعلمين فكما هو متعارف عليه هنالك تطبيق عملي يتدرب فيه الطالب على مهارات التدريس من خلال برنامج التربية العملية الذي يضم مراحل المشاهدة والتطبيق الموزع والتطبيق المكثف وفي هذا المجال يتم اختبار الطالب بواسطة اختبارات الأداء

الأسس التي تتبع في تحضير اختبارات الأداء

تعتبر الاختبارات الأدائية جزءا من الاختبارات النهائية التي يطلب من الطلاب اجتيازها للنجاح والتخرج. وفيما يلي تحديد للأسس أو الشروط التي يجب مراعاتها قبل القيام بالاختبار الأدائي:

1- تحديد أهداف الاختبار الأدائي: يجب على الفاحص أن يجيب على عدد من الأسئلة مثل: ما الذي سوف ألاحظه؟ وما هو معيار النجاح؟ كيف ستتم تأدية هذا السلوك؟

2- تأكد الفاحص من استعداد الطالب للامتحان وانه قد أتم تدريبه بما يكفي لأداء الاختبار.

3- تحديد المكان والزمان اللذين سيتم فيهما الاختبار الأدائي.

4- تقويم الطالب من قبل لجنة يتم الاتفاق فيما بين أعضائها حول أداء الطالب.

5- تقويم الطالب بناء على بطاقة ملاحظة تتم من خلالها مراقبة مدى تحقق الأمور الواردة في بطاقة الملاحظة.

6- تدخل اللجنة الفاحصة في حالة وجود خطر حقيقي، ومن الأمثلة على ذلك وجود خطأ في تجربة كيميائية قد يؤدي إلى انفجار أو غيره.

7- اعتماد الاختبار الأدائي على وحدات متساوية، وهذا يعني اعتماد الاختبار على المساواة بين المهارة وعدد الدرجات من جهة، واعتماد الاختبار على أرقام صحيحة بدون كسور من جهة أخرى.

8- تحليل المهارة التي يريد الفاحص اختبار الطالب بها إلى مكوناتها لتسهيل قياسها.

أهداف اختبارات الأداء:

للاختبارات الأدائية عدد من الأهداف أهمها:

1- قياس مدى فهم الدراسة النظرية المرتبطة بالعلوم الطبيعية، كالكيمياء والفيزياء وغيرها.

2- قياس مدى التحصيل الذي وصل إليه الطالب في مواد التدريب العملي في المدارس الصناعية وما شابهها.

3- قياس بعض المهارات مثل مهارة الطباعة والسكرتارية والموسيقى وغيرها.

4- التنبؤ بنجاح الطالب مستقبلا في مهنة معينة.

5- تشخيص جوانب الضعف في أداء الفرد لمهارة معينة وبالتالي وضع أسلوب علاجي لذلك.

طرائق تصحيح الاختبارات الأدائية وتفسير نتائجها

يمكن تفسير نتائج الاختبارات الأدائية بعدة طرق هي:

1- استخدام بطاقة للاختبار يحدد عليها جوانب المهارة وإعطاء أرقام أو درجات يتم تقدير درجة الطالب على الاختبار الأدائي من خلالها.

2- استخدام الاختبارات المحكية، بحيث يحدد المعلم محكاً معيناً يقارن أداء الطالب بناءً عليه لتقويم هذا الأداء.

3- استخدام الاختبارات المرجعية المحك، وبالتالي يمكن أن يستخدم اختبار مقنن معد لذلك كاختبار أداء أو أن يقارن أداء الطالب بغيره.

4- استخدام الدرجة المعيارية والمئينات.

الفصل التاسع

اختبارات الذكاء

يحتوي الفصل على:

- ماهية الذكاء
- نبذة تاريخية عن اختبارات الذكاء
- طبيعة اختبارات الذكاء
- عماذا تخبرنا اختبارات الذكاء وبم تتنبأ؟
- وحدات الذكاء
- طبيعة الذكاء المقاس
- ثبات اختبارات الذكاء
- الوراثة
- التشابه في نسب ذكاء ذوي القربى
- التشابه الوراثي أو البيئي
- مفهوم الموروث
- ما الذي نعنيه بالموروث
- نماذج من اختبارات الذكاء

ماهية الذكاء

لمواجهة أول مشكلة لنا في سبيل فهم هذا المفهوم المعقد "الذكاء" فإننا نحتاج فقط إلى محاولة لتعريفه. الذكاء كلمة تظهر لنا جميعا أننا نفهمها ونستخدمها بشيء من التوافق أو التشابه إلا أن قليلاً منا، سواء على مستوى الشخص العادي أو العالم المتخصص، يمكن أن يتفق على تعريف لفظي لهذا التعبير وبمرور الزمن بدأ العلماء يتكلمون عن الذكاء كقدرة على:

- التفكير أو المنطق المجرد
- التعلم أو الاستفادة من الخبرة
- تكيف الشخص مع بيئته
- حل المشاكل
- تجميع المعرفة والمعلومات... الخ

وعلى عكس الشخص غير المتخصص، ولحاجة العالم (المتخصص) إلى تعريف المفهوم بدلالة الإجراءات أو القياسات، فإن بعض العلماء اقترحوا تعريفاً سهلاً له (أي الذكاء) على أساس إنه: "ما تقيسه اختبارات الذكاء".

مثل هذا التعريف يحل المشكلة ولكنه يخضع للسؤال "فالتعريف الإجرائي يكون مفيداً متى ما اتفق الجميع على أن هذا الإجراء أو القياس يعكس المفهوم المعرف بتمام أو صدق. إن مثل هذه الإستراتيجية في التعريف تعمل بجودة بالغة بالنسبة لمفهوم "الوقت" عندما يقاس بساعة مضبوطة، أو مفهوم "الوزن" عندما يحدد من قبل ميزان دقيق محكم، غير إنه حتى بالنسبة لهذين المفهومين (الوقت والزمن) والإجراءات المستخدمة فإن هناك معياراً موحداً مقبولاً على المستوى العالمي يمكن له أن يخبرنا عن مدى دقة تلك الساعة أو ذلك الميزان. وإذا ما كان الذكاء هو ما تقيسه اختبارات الذكاء، ولاشيء اقل من ذلك أو أكثر فكيف تتأتى لنا معرفة ما إذا كنا نمتلك اختباراً جيداً للذكاء؟ ثم مادام هناك عدة اختبارات للذكاء ولا تعطى جميعها ذات الدرجة من الذكاء لذات للفرد: فهل يمكن لأحدنا

أن يفترض وجود عدة أنواع من الذكاء بقدر عدد اختباراتها؟ فضلاً عن ذلك فإن معظمنا يشعر أن هناك أنماطاً من السلوك الذكائي أكثر منه أنواعاً من الذكاء فيما تعرضه هذه الاختبارات أي اختبارات الذكاء وهنا يظهر إننا ندخل مازقا أو طريقاً مسدوداً. حتى لو إننا لا نستطيع الاتفاق على تعريف دقيق لكلمة (الذكاء) فإن العلماء اتفقوا عموماً على أن الذكاء ليس كينونة (شيئاً موجودا Entity) أو وجودا يمكن أن يتملكه الناس كما يفعلون في الدم أو الدماغ. إن تعبير الذكاء مفهوم قريب (مماثل) لمفاهيم أخرى كالكسل، الدقة، البهجة. ومثلها تماماً فإن الذكاء يعني ضمناً مجموعة من الأنماط السلوكية التي نرغب في وصفها ذكاء أكثر دقة فإننا يجب أن نرجع (Refer) إلى الأنماط السلوكية للفرد وليس إلى ذكائه. فضلاً عن ذلك فا ن مجموعة الأنماط السلوكية غير محدودة أساساً وقد تحتوي على عدة إنجازات متباينة:

- كحل مسألة رياضية معقدة.
- أو تصميم جسر.
- أو رسم مبدع للوحات يجد فيها الآخرون متعة ولذة.
- صياد ماهر.
- النجاة من جزيرة غير مأهولة.

إن حقيقة كون تصنيف أنماط السلوك الذكائي من السعة والتباين بمكان، قد حدا ببعض علماء النفس إلى أن يندبوا تعبير الذكاء بأن لا معنى له إطلاقاً ويجب أن ينحى كلية، حيث أشاروا إلى الخطورة من إعطاء كلمة واحدة لمثل هذه المجموعة المتعددة من الجوانب والأوجه من الأنماط السلوكية. وقديما حذرنا John steward mill من ضعف إنساني معين بقوله "دائما ما تكون النزعة قوية للاعتقاد بأن أي تقبل لأسم لابد أن يكون لشيء موجود ويمتلك وجوداً مستقلاً (خاصا به). وحين لا تكون هناك إجابة وجودية حقيقية للاسم المراد إيجاده، فإن الناس بدلا من افتراضهم عدم وجوده، يتخيلونه كشيء كان غامضاً مبهماً وعلى نحو غريب. وعلى كل، ومهما كان الذكاء فهو ليس شيئاً موجودا يمكن أن نحمله معنا، إنما هو

يرجع إلى عدد غير محدود من الأنماط السلوكية التي اخترناها لتصنف لنا الذكاء. إن مفهوم الذكاء يماثل مفهوم البسالة الرياضية، فهي ليست شيئاً يظهر على الشخص اسمه البسالة الرياضية إنما يعرف هذا المفهوم بدلالة إنجاز الشخص الكامن أو الفعلي عند أدائه لأنماط سلوكية معينة نسميها الرياضة. أكثر من ذلك فإن هناك عددا لا يمكن حسابه تقريباً من السلوكيات الرياضية المختلفة وفي الوقت الذي توجد فيه نزعة عند الشخص الجيد في نوع معين من الرياضة أن يكون جيداً في ألعاب أخرى كذلك فإن هناك حدوداً نقف عندها عندما يراد تعميم مثل هذه القدرة الرياضية. إن الشخص الطويل الخفيف الحركة، القوي، المتسق، ربما يتفوق في كرة السلة، كرة القدم، أو القفز العالي، إلا إنه لا يمتلك خبرة معينة في الرماية، تنس الطاولة، أو في ركض المارثون.

قياس الذكاء

إذا ما كان تعريف الذكاء من الصعوبة بمكان أليس من البلاهة أساسا محاولة قياسه؟ ربما يبدو كذلك، إلا أن العلماء في علوم معرفية أخرى يقيسون مفاهيم شبيهة (بمفهوم الذكاء) ليس لها تحديد وجودي في هذا العالم. ولنأخذ مفهوم الجاذبية على سبيل المثال: مراجع الفيزياء لا تعرف لنا الجاذبية، ومادام الفيزيائيون غير متفقين على الطبيعة الأساسية لنظرية: الجذب والمجال، محاور الفراغ الوقت، الكون الأحدب ...الخ فإن ما ترجع إليه مراجع الفيزياء إنما هي أمثلة يعني بها المفهوم (كسقوط التفاحة مثلا) ثم تذهب بعد ذلك لوصف الكيفية التي يمكن لأحدنا فيها قياس الجاذبية. ولكن إذا ما قسنا الذكاء كشيء من غير كينونة محددة، إلا يقودنا ذلك إلى جدل بيزنطي (دائري) من النوع التالي: (شخص ما درجته جيدة في الاختبار بسبب ذكائه، وهو ذكي لان درجته في الاختبار كانت جيدة). مرة أخرى فإن العبارة هذه تبدو من غير فائدة أو نفع أيضاً إلا أن العلماء في فروع المعرفة الأخرى يعملون ذات الشيء أيضاً: (نحن نقيس الجاذبية بملاحظة سقوط التفاحة إلى الأرض، ثم ننقلب بعد ذلك لتوضيح سقوط التفاحة عن طريق مفهوم

الجاذبية) والحقيقة، فإن كلا من المقياس (الاختبار) والمفهوم يصبح مفيداً عندما يمكن تجميع الكثير من الأمثلة التي تعود إلى نتائج الاختبار ولنأخذ أمثلة أخرى على ذلك:

- إذا ما اتبعت أقمار (توابع) الكواكب ذات القوانين التي تسقط بها التفاحة
- وإذا ما سلك الضوء (المكون من جزيئات) نهجاً موافقاً لمفهوم الجاذبية كذلك ...الخ

فإن كلا من القياس ومفهوم الجاذبية سيكون نافعا ولحد ما صادقا كذلك تماما فيما يخص الذكاء، إذا ما تنبأ الاختبار بعدد من الأنماط السلوكية العقلية المترابطة (المتداخلة) فإنه يمكننا افتراض وجود عضو عام (مشترك) أو مفهوم يؤكد التشابه ما بين هذه الأنماط السلوكية بين الأفراد. ربما يمكننا تسمية هذا المفهوم (ذكاء) ونقبل الاختبار كمقياس له.

ومع عدم قدرتنا لتعريف تعبير الذكاء بدقة، ومع مضي الوقت امتلك الناس بدلا منه، شعورا واضحا (محددا) لما يمكن (للذكاء الحقيقي) أن يكون عليه.

الذكاء الحقيقي سمة موحدة تؤكد إنجاز الفرد في كل المهمات التي نعتبرها عقلية أو ذهنية. وهو ثابت عبر حياة الفرد، وهو نوع من التنظيم (النزعة) الأساسي المستثنى (نسبيا) من الخبرة وفرص التعليم، لذا فإن التعليم والتشجيع يساعدان على التعبير عنه وترجمته أي (الذكاء) إلى (سلوك الذكائي) ومن هنا اعتقد الناس أن (الذكاء الحقيقي) هو غير مقيدا بثقافة ما Culture-Free ويتوزع بالتساوي على سكان العالم والذين يختلفون بسهولة من شخص إلى آخر في طريقة استخدام (أو التعبير عن) ذكائهم.

إن واحدا من عوامل الجدل المعاصر حول الذكاء، اختباره، الوراثة، الجنس (العنصر Race)... الخ يخص مفهوم الذكاء الحقيقي في يد واختبارات نسب الذكاء (IQ- Tests) في يد أخرى.

تاريخ اختبارات الذكاء

طلبت وزارة التعليم الفرنسية من الفريد بينيه ALFRED BINET تطوير وسيلة تقديرية يمكن عن طريقها استبيان أو اكتشاف مَن مِن أطفال المدرسة يمكنه النجاح في النظام الفرنسي المدرسي العادي، ومن منهم يحتاج إلى تعليم خاص. لاحظ أن وزارة التعليم لم تسال بينيه تطوير اختبار الذكاء انما أرادوا وسيلة تقديرية لهدف عملي محض.

إن "اكتشاف الطلبة الذين لا يمكنهم الاستفادة من البيئة المدرسية العادية مسبقاً، أي قبل دخولهم المدرسة". ولم يكن في ذهن وزارة التعليم ولا في ذهن بينيه إيجاد أو إنشاء اختبار للذكاء العام كما نعتبرها (اختبارات الذكاء المعاصرة) نحن الآن. أنتج بينيه بمساعد الطبيب Pro. Seimon الاختبار الأول في عام 1905، ثم تلا ذلك اقتراح William Stern بأخذ نسبة العمر العقلي Mental age للشخص كما يحدده الاختبار، وتقسيم على ذلك عمره الزمني Chronological Age لنخرج بما نسميه الآن (نسبة الذكاء Intelligence Quotient) أو درجة نسبة الذكاء IQ-Score. إن هذه الدرجة المفردة استطاعت أن تحدد بسهولة بالغة من يجب أن يذهب إلى المدرسة ومن لا يذهب.

بعد النجاح الأولي للاختبار نقح بينيه اختباره في عام 1908 ثم أعاد تنقيحه ثانية في عام 1911 ليكون أكثر ملائمة للتمييز بين مستويات الإنجاز في المدارس الاعتيادية بدلا من استبيان المتخلفين المعتدلين (Moderate Retardation) أي غير المتطرفين في تخلفهم فقط

وفي عام 1916 قام لويس تيرمان Lewis M. Terman بنشر اختبار ستانفورد – بينيه Stanford- Binet Test للذكاء وهو صيغة معدلة لاختبار بينيه لاستخدامه في الولايات المتحدة الأمريكية ثم قام تيرمان بعد ذلك بتنقيح طبعته في عام 1937 في شكل صيغتين متكافئتين جوهريا سماهما (L & M) ثم جمع أفضل السمات

(Features) في هاتين الصيغتين بطبعته الجديدة لعام 1960 ليضعها في صيغة واحدة (كما بدأ الاختبار أول مرة) تحت اسم (L-M).

استحوذت طبعة تيرمان لاختبار بينيه على مجال القياس أو التقدير العقلي لعشرين عاما قبل أن يخرج (ديفيد ويكسلر David Wechsler) مقياسا آخر للذكاء هو (مقياس ويكسلر - بيليفو للذكاء) Wechsler- Bellevue Intelligence Scale والذي عرف فيما بعد بمقياس وكسلر لذكاء الراشدين WALS: Wechsler Adult Intelligence Scale اختلف مقياس وكسلر عن اختبار ستانفورد - بينيه بما تضمنه من بطاريتين (مجموعتين) من الاختبارات الفرعية لكل مستوى من المستويات العمرية:

بطارية الاختبارات اللفظية Verbal Tests، وتحتوي على اختبارات

- المعلومات Information
- الإدراك(الفهم) Comprehension
- الأرقام الصاعدة والنازلة (العدد الطردي والعكسي) Digits forward & Backward
- المتشابهات Similarities
- المفردات Vocabulary

بطارية اختبارات الانجاز Performance tests، وتحتوي على اختبارات:

- تكملة الصور Picture completion
- ترتيب الصور Picture Arrangement
- تركيب (تجميع) الأشياء Object Assembly
- تشكيل القطع (المكعبات) Block design
- الرقم والرمز Digit symbol

وبالإضافة إلى اختبار وكسلر للراشدين فهناك طبعة خاصة بالأطفال تسمى

(WISC) أي مقياس وكسلر لذكاء الأطفالIntelligence Scale for children وطبعة أخري كذلك لذات الاختبار خاصة بأطفال ما قبل المدرسة الابتدائية والصفوف الأولية تسمى (WIPPSI) أي: مقياس وكسلر لذكاء أطفال ما قبل المدرسة والصفوف الأولية Wechsler preschool and primary scale of intelligence.

إن اختبارات ستانفورد - بينيه ووكسلر هي التي تستخدم أساساً كاختبارات لنسب الذكاء الفردية أي تعطى للأفراد المفحوصين فرداً فرداً كل على حدة) في الوقت الحاضر. كما أن هناك عدة اختبارات أخرى تسمى نفسها اختبارات ذكاء، بعضها قصير نسبياً وبعضها يمكن تطبيقها على شكل أفراد أو مجموعات، ولكن معظمها لم تحظ بالقبول أو الانتباه الذي أعطي لاختبارات بينيه ووكسلر. إن من المهم بالنسبة لنا ملاحظة أن هذه الاختبارات لم تشتق من محاولة قياس مفهوم الذكاء فقد تعامل بينيه (كما ذكر سابقاً) مع مشكلة عملية صرفة لكشف أو استبيان الأطفال الذين يتوقع فشلهم في المدرسة. وحتى في كتاباته الأخرى لم يتناول بينيه ضمن آرائه النظرية مكونات أو عناصر القدرة العقلية أو الذكاء. لذا فإن من الواضح إنه لم يؤيد أو يقر " القدرة المفردة الفطرية العامة: Single Inborn General Ability. وعلى كل حال فالحقيقة، أن ماهية فكرة بينيه عن الذكاء ليست ذات أهمية مادام اختباره (بينيه) نجح نجاحا باهرا ومع تنقيحه لم يصبح قادرا على التمييز بين النجاح والفشل المدرسي فقط، إنما تعدى ذلك إلى التمييز بين مستويات الانجاز المدرسي العادي. ومادام الشخص يحتاج بوضوح إلى الذكاء لكي ينجز جيدا في المدرسة وان اختبار بينيه يتنبأ بالنجاح المدرسي فإنه ليس صعبا افتراض أن نسبة الذكاء يمكن أن تكون مقياسا لمفهوم الذكاء وللقوة التي أصبح بها الاعتقاد قويا بأن درجة (نسبة الذكاء) هي ذكاء الفرد، فإن معظم الاختبارات العقلية الأخرى قد حكم عليها من خلال علاقتها (مرجعيتها) لاختبار ستانفورد -بينيه ثم اختبار وكسلر بعد ذلك، بسبب افتراضهم بالطبع جودة هذه الاختبارات كمقياس للذكاء، ومادام المعتقد الشائع أن الذكاء خاصية ثابتة نسبيا عند الأفراد، فإن اختباراً واحدا

لتحديد نسبة الذكاء (استنادا إلى هذا المعتقد) كان يعتبر كافيا وليس هناك من ضرورة لإعادة هذا التحديد أو التقدير. فضلا عن ذلك فإن الاعتقاد بأن الاختبارات تقيس الذكاء في جميع صيغه وعروضه الممكنة كان قويا جدا ولدرجة إنه متى ما أريد انتقاء أشخاص لمهنة أو وظيفة معينة (يبدو أنها تستلزم ذكاء) استخدمت اختبارات نسب الذكاء كأداة لغربلتهم بدون تردد من غير حتى أن يسأل هل أن نسب الذكاء هذه على علاقة حقيقية بنجاحه في تلك الوظيفة ام لا؟ أن الوظيفة تحتاج إلى الذكاء والاختبارات تقيس الذكاء فما الذي كانوا بحاجة إليه من ذلك؟

طبيعة اختبارات نسب الذكاء

رجوعا إلى اختباري ستانفورد - بينيه ووكسلر فإن هناك فروقا معينة مابين اختبارات بينيه ووكسلر.

مثلا

تمتلك اختبارات وكسلر فقرات اختبارات غير لفظية "للانجاز" أكثر من اختبار بينيه. إلا أنهما (أي اختباري بينيه ووكسلر) يتشابهان كثيرا في السمات العامة حيث تنزع الأسئلة عادة إلى اللفظية في طبيعتها، وغالبا ما تتطلب بعض التفكير أو حل لمسالة تتعلق بسمة (أو خاصية) مجردة وبالذات تلك الفقرات الاختبارية المخصصة للأطفال الكبار أو الراشدين.

مثلا

- ربما يسأل الفرد تعريف كلمات معينة.
- أو يصف نوع البلاهة في عبارة سخيفة أو غير معقولة.
- أو يسأل ليخبرنا كيف يتشابه شيئان أو ثلاثة مع بعضهما (البرتقالة والموزة مثلا)

- أو مسائل لفظية (مثلا: إذا كنت ذاهبا باتجاه الغرب، وعملت استدارتين إلى اليمين، ثم استدارت بعد ذلك إلى اليسار، فبأي اتجاه أنت تسير الآن)
- أو إدراك فهم لمواقف معينة (لماذا نمتلك السيارات مثلا؟)
- أو إجراء صرف للعملة عند شراء شيء.
- أو تحديد (تعيين) كم سيأخذ شخص من الوقت لقطع مسافة 12 ميلا بمتوسط سرعة 4 ميل لكل ساعة.
- أو استذكار لسلسلة أرقام طردا أو عكسا
- أو حل بعض المهمات المتعلقة بالملاحظة أو الإدراك الحيزي (الفراغي) Spatial Perceptual (مثلا عمل تصميم ملون من مجموعة 16 قطعة (مكعبا).

هذه الفقرات درجت بحسب العمر أي إنه خلال بناء الاختبار أعطيت الأسئلة (الفقرات الاختبارية) إلى عينة كبيرة من الأطفال بمختلف الأعمار ولوحظ بأي عمر ينزع الأطفال إلى القدرة على الإجابة عن تلك الفقرة صحيحا ثم رتبت هذه الفقرات الاختبارية بحسب العمر، ولكل منها - أي الفقرة الاختبارية - قيمة تمثل عدة أشهر من العمر العقلي.

إن الاختبار ونظام التدرج صمما بحيث أن طفل السادسة من العمر (72 شهرا) سوف يجيب صحيحا عن ما يكفي من الأسئلة لكي يحصل على عمر عقلي قدره (72 شهرا) أو عندما يقسم العمر العقلي (M A) على العمر الزمني (CA) ونضرب ب(100) فإن الناتج لذلك سيكون:

$$IQ = MA/CA \times 100 = 72/72 \times 100 = 100$$

عماذا تخبرنا اختبارات الذكاء؟ وبم تتنبأ؟

لا قيمة للاختبار ما لم تكن الدرجات التي تقدمها على علاقة بما قصد أو هدف الاختبار أن يقيسه. ففي حالة اختبار بينيه فقد قصد الاختبار قياس النجاح المدرسي

(أو التنبؤ به) وليس الذكاء لذا فإن صدق اختبار بينيه يعتمد على مدى تنبؤه بالنجاح المدرسي. ليس هناك من شك ان الدرجة في اختبار بينيه أو ويكسلر لها علاقة كبيرة جدا بالإنجاز الأكاديمي الحالي للطفل ولحد ما بمستقبل نجاحه المدرسي.

وللحقيقة فإن اختبارات نسب الذكاء هي المتنبئ الأفضل الفريد بالنجاح المدرسي في الولايات المتحدة الأمريكية، وهذه الحقيقة تصدق على مجتمع الأغلبية أو الأقليات على حد سواء، اليوم كما كانت بالأمس.

وفي هذا المعنى أو المفهوم فإن اختبار نسبة الذكاء اختبار جيد أو صادق، ولحد ما في الحقيقة فأنها قد تكون أفضل اختبارات وجدت لحد ألان ولكن علينا أن نتعجل إضافة أن الاختبار غير صادق أو إنه ليس الأفضل بصورة مجردة كذلك. فالاختبار صادق لغرض معين فقط. فهي -اختبارات نسب الذكاء - صادقة للنجاح المدرسي، إلا أنها غير ذات جدوى للتنبؤ بالبسالة الرياضية أو سرعة الطباعة وفي الحقيقة، فهي حتى غير قادرة على التنبؤ بكل أنماط الإنجاز المدرسي على حد سواء.

مثال

- إن اختبارات الذكاء على علاقة جيدة جدا باستيعاب القراءة Reading Comprehension.
- وعلى علاقة حسنة باستخدام الانكليزية English Usage والتاريخ.
- ولكنها بالأحرى اقل دقة أو ضبطا مع سرعة القراءة Reading Speed والهندسة.

والآن: ماذا بشأن الإنجازات الأخرى التي نصفها عادة كذلك؟ مثلا. هل تتنبأ اختبارات نسب الذكاء بالمكانة الاجتماعية Social Status الجواب: نعم ولحد ما ولكن ليس إلى الدرجة التي يتنبأ بها التحصيل الأكاديمي. أكثر من ذلك فإننا نعلم أن النجاح المهني Occupational يمكن أن يحصل عليه الفرد لحد ما من خلال تعليمه

سواء لما تعلمه الفرد في المدرسة أو من خلال الفرص المفتوحة للأفراد الذين يتسلقون السلم الأكاديمي. وفعلا نحن قادرين على التنبؤ بالنجاح المهني من خلال معرفتنا لمستوى الفرد التعليمي كاستطاعتنا له من خلال معرفة نسبة ذكائه. بل أن نسب الذكاء على علاقة اقل بمستوى "الدخل Income " بالرغم من الحقيقة التي غالبا ما نفترضها من أن هناك علاقة قوية مابين الذكاء والغنى: (فهو دليل التحدي: إذا ما كنت ذكيا جدا فلماذا أنت غير غني؟).

في الواقع، فإن العلاقة مابين نسب الذكاء والدخل منخفضة إلى حد ما، لأننا عندما ندفع مرتبا جيدا للفرد ليس فقط لان الوظيفة تتطلب مهارة فذة (نادرة) وإنما عندما تكون الوظيفة غير مرغوبة أو خطرة كذلك.

إن من السهولة القفز إلى استنتاج أن اختبارات نسب الذكاء مفيدة فقط كمتنبئات بالنجاح الأكاديمي، إلا ان هذا ليس صحيحا. لان ذات المهارات أو المهارات ذات العلاقة المستلزمة للنجاح في المدرسة هي ضرورية كذلك لعدد كبير من الفعاليات التي تبدو لنا أنها غير أكاديمية.

مثلا قد تأخذ عند معظمنا بضع دقائق فقط من الحديث عندما يتكلم عن الفرق مابين شخص يحصل على (85) و آخر على (115) في اختبار نسبة الذكاء.

- إلا أن نوع الوظيفة المناسبة لكل منهما ربما يكون مختلفا.
- كذلك رغباتهم وفعالياتهم وقت الفراغ يحتمل أن يكون غير متشابهة تماما.
- قد يقرؤون مجلات مختلفة ويشاهدون برامج تلفزيونية مختلفة.

أكثر من هذا أهمية، فلقد شوهد أن نسب الذكاء على علاقة بنوعية الرعاية الغذائية والسلوكية التي تقدمها الأمهات لأطفالهم، وهي عوامل قد تكون ذات تطبيقات أو استخدامات واسعة الانتشار لهؤلاء الأطفال. وعندما يفكر احدنا بالعدد الكبير من الأنماط السلوكية، المهارات، الفرص، الرغبات ذات العلاقة لقدرة الفرد للنجاح في المدرسة، فإنه يمكن لأحدنا أن يفهم لماذا تكون علاقة

درجات متفاوتة لكثير من الأنماط السلوكية المختلفة والإنجازات فيما بعد (ما وراء) مجال أو حقل النجاح الأكاديمي.

نخلص إلى:

أن اختبارات نسب الذكاء المعاصرة تتنبأ بالنجاح المدرسي في الولايات المتحدة الأمريكي شأنها شان أية أداة تقدير assessment مفردة أخرى، وهي جميعا على ذات القدرة من التنبؤ لأطفال الأغلبية أو الأقليات على حد سواء. وعلى كل، فإنه كلما كانت المحكات (المعايير) على علاقة أقل فاقل مع الإنجاز المدرسي فإن اختبارات نسب الذكاء تكون اقل كفاءة أو مهارة في التنبؤ. كذلك فإن اختبارات نسب الذكاء على علاقة بالنجاح في أنواع كثيرة من المهمات والأنماط السلوكية، وبالطبع فإن لها حدودها.

وحدات الذكاء

إن من المهم ملاحظة الفرق مابين أن نقول أن اختبار نسبة الذكاء صادق للتحصيل المدرسي من جهة وان نؤكد أنها تخبرنا كم يمتلكون من الذكاء من جهة أخرى. إن نقطة نسبة الذكاء –ببساطة– هي ليس وحدة (الذكاء) كما هو المعنى في أن يكون (الباوند) وحدة (الوزن) أو (الانج) وحدة (للطول). إن الشخص الذي يسجل (160) كنسبة لذكائه في اختبار للذكاء هو ليس على درجة الضعف في ذكائه من شخص يسجل (80)، كما إنه لا يمتلك - ذات الوقت - ثمانين وحدة أكثر من القوة العقلية للثاني. إن درجة نسبة الذكاء - ببساطة - عبارة عن الإنجاز النسبي

Relative performance على مجموعة من الأسئلة ترجع لأفراد من عمر مماثل وان الأعلى درجة فيهم يتوقع منه أن ينجز في المدرسة بصورة أفضل.

طبيعة الذكاء المقاس

سبق وان أوضحنا أن معظم الناس يعتقدون أن الذكاء الحقيقي سمة مركزية (وحدية Unitary) تحدد إنجازاتنا في جميع الأنماط السلوكية التي تدعوها عقلية أو ذهنية وهو ثابت عبر عمر الإنسان وغير متحيز لثقافة معينة، ويتوزع بالتساوي خلال المجموعات الثقافية.

ولسبب ما شعر الناس أن الذكاء الحقيقي يجب أن يكون مفردا وموحدا، وان الأنماط المختلفة للسلوك العقلي التي تلاحظها عند الأفراد يجب أن تنشأ وتنبع كلها من لب أو جوهر واحد هو الذكاء. وللقوة التي كان عليها مثل هذا الاعتقاد فإن بعض الأفراد يجادلون بأنه لو كان الذكاء سمة غير وحدية، فبطريقة أو بأخرى فإن المفهوم الكلي سوف يتحطم. فمتى ما كان هناك عدة أنواع من الذكاء فإنه لن يكون حينذاك ذكاء على الإطلاق. وللشأن والأهمية فلابد من التنويه ان مثل هذه المشكلة قد ظهرت في علوم أخرى: فقبل حوالي (150) عاما أثير ذات النمط من الجدل حول ما إذا كانت الكهربائية نوعا واحدا من القوة (قوة وحدية) أو يمكن تجزئتها إلى عدة قوى مثل الكهربائية الاستاتيكية (الساكنة Static) أو الكهربائية الفولتية (المتحركة Voltaic). فقد وجد أن الكهربائية الفولتية تظهر ضمن المعدن (الفلز) وتعتمد على سمكه - أي المعدن - بينما يمكن أن تنتقل الكهربائية الحرة حتى على اخف المعادن سمكاً وبالطبع فإن مفهوم الكهربائية يستخدم بإفراط هذا اليوم بالرغم من أنها ليست وحدية. كما أن حقيقة كون الذرة تحتوي (بالفعل) على عدة أنواع من الجسيمات الأولية Elementary Particles لم تحطم استخدام مفهوم الذرة. لذا فإن مفهوم الذكاء سوف لن يتحطم بالضرورة إذا وجد إنه -أي الذكاء- يمتلك أكثر من صيغة أو إنه يغير خصائصه بتقدم العمر.

والآن أين تأصلت فكرة الذكاء الوحدي (المركزي) العام؟ بعد أن وجدت اختبارات نسب الذكاء واختبارات لعدة مهارات عقلية أكثر نوعية ولوحظ أن كل اختبارات الإنجاز العقلي تقريبا ترى درجة ما من العلاقة الموجبة بعضها للبعض الآخر.

معنى ذلك أن الشخص الذي يسجل درجة عالية نسبيا في احدها (أي الاختبارات العقلية) يحتمل إنه سيسجل عاليا في اختبار عقلي آخر بالرغم من أن الإنجازات اللازمة لهذين الاختبارين العقليين تظهر مختلفة جدا. فقد افترض إنه مادام الناس يسجلون أما عاليا أو واطئا بعدة مهمات عقلية فإن سلوكهم في هذه التقديرات المتنوعة لابد أن تتوسط بذكاء مفرد مركزي. فإذا ما أعطى أحد الأفراد عددا من الاختبارات العقلية، فإن متوسط درجاته في هذه الاختبارات المتباينة سيكون مقياسا لذكائه العام والذي سمى "بالقدرة العقلية العامة Genaral mental ability من قبل شارلز سبيرمان Charles spearman أحد الأعلام الرواد في تحليل الأنماط (الأنواع) المختلفة من الاختبارات العقلية حيث أطلق عليها الرمز (G) ولكن حتى سبيرمان الذي تزعم فكرة الوحدية هذه، أي الذكاء العام وجد أن هذه الاختبارات تختلف في مدى علاقة هذه الدرجات بدرجات عقلية أخرى وان الانجاز في بعض المهمات العقلية كان على علاقة قليلة (ضعيفة) بالقدرة العقلية العامة (G) وبالتالي فلابد من وجود عوامل أخرى تحدد انجاز الفرد في الاختبار المعطي والتي تخص تلك المهمة بالعين.

من هنا فإن الإدراك الأولي الأكثر بروزاً للذكاء المقاس كان وجود ذكاء عام قوي وحدي G تلحق به - بتقدير اقل أهمية - عوامل محددة تعود إلى أنواع معينة من المهمات و المهارات. ما حدث بعد ذلك في تاريخ الاختبار العقلي كان عبارة عن محاولات هذه لتحديد ماذا كان (G) موجودا وان هذه القدرات الخاصة في الحقيقة لم تكن ذات أهمية اكبر من الفكرة الأصلية. لقد أكد سبيرمان على أهمية القدرة العقلية العامة الوحدوية، من خلال اعتماده على سلسلة من التجارب وتهدف طريقته إلى اختيار طائفة من الأفراد من خلال مجموعة من الاختبارات التي تقيس نواحي النشاط العقلي المعرفي كما تبدو في القدرات العقلية المختلفة وفي التحصيل المدرسي. ومن ثم تم إيجاد معاملات الارتباط بين هذه المقاييس وترتيب هذه المعاملات ترتيباً بالنسبة لقيمها العددية في جدول أو مصفوفة ودراسة الخواص

الإحصائية النفسية لهذا الجدول أو لتلك المصفوفة للتأكد من أنها تدل على وجود عامل واحد والكشف عن مدى احتواء كل اختبار من تلك الاختبارات على هذا العامل العام.

وتابع سبيرمان دراساته في نظرية الوحدة العقلية التي تؤدي منطقياً إلى ما يسميه "الترتيب الهرمي لأنماط الذكاء النوعي" وكانت طريقته في صورتها الأولى تقريبية، إلا أن سبيرمان لجأ إلى طريقة أخرى هي حساب معامل الارتباط بين الأعمدة للوصول إلى تحديد دقيق للترتيب الهرمي وصولا إلى طريقة أكثر دقة تسمى بـ "محك الفروق الرباعية" وهي التي اقترنت باسمه، حيث تكمن فيها بدايات منهج تحليل المحتوى. لقد بنيت العديد من الاختبارات على وفق نظرية العاملين منها، اختبارات سبيرمان الحسية للذكاء، واختبارات المصفوفات المتتابعة لرافن، واختبار الذكاء المتحرر من اثر الثقافة لكاتل.

إلا أن هذه النظرية لم تسلم من النقد مما دعا سبيرمان وتلامذته للتراجع عن الترتيب الهرمي، والاعتراف بوجود عامل مشترك فضلا عن العامل العام في بعض الاختبارات. وإزاء ذلك قدم هولزنجر طريقة العوامل المزدوجة تلافياً لطريقة سبيرمان. ومن ابرز النقد ما يلي:

- إن نموذج العاملين الذي قدمه سبيرمان ليس إلا تفسيرا واحداً محتملاً للنتائج الإحصائية التي توصل إليها.
- صغر حجم العينة التي أقام سبيرمان نظريته عليها إلى جانب قلة الاختبارات المستخدمة.
- يختلف العامل العام من تجربة لأخرى وفقاً لنوع الاختبارات وصفات الأفراد.
- كشف العيد من الباحثين وجود عوامل طائفية غير العامل العام الذي ذكره سبيرمان.

ويرى هولزنجر إنه يمكن الحصول على العامل العام أو أكثر من العوامل الطائفية.

فتباين الاختبار يشمل (العامل العام والعوامل الطائفية والعوامل الخاصة) وان طريقة هولزنجر لاتركز على عامل واحد وإنما تسمح للاختبارات التي لا ينطبق عليها محك الفروق الرباعية أن تظل في مصفوفة الارتباط وذلك بتفسير جزء من تباينها المشترك في ضوء العوامل الطائفية.

أما ثورندايك فقد انتقد نظرية العامل العام لسبيرمان معتقدا أن الاختبارات المستخدمة من قبل سبيرمان سبباً في الوصول إلى هذا العامل إلى جانب قلتها. وأكد أن الذكاء يتكون من عدد كبير من العناصر أو العوامل المنفصلة، حيث أن كل أداء عقلي عبارة عن عنصر منفصل ومستقل إلى حد ما عن بقية العناصر الأخرى. ويعود استخدام (المثير -الاستجابة) في تفسير الذكاء إلى ثورندايك لعام 1911 حيث يرى أن التعلم هو تكوين ارتباطات أو تقوية الارتباطات بين المثير والاستجابة، ولذلك فينظر ثورندايك للذكاء بصورة رئيسية على أساس القدرة على التعلم. إلا أن كانيه 1970 يرى بأن هناك ثمانية أنواع من التعلم والتي تختلف بالنوعية والكمية من حيث الروابط بين المثير والاستجابة.

ويبدو أن ثورندايك وجه انتقاداته لنظرية سبيرمان حيث رفض فكرة عمومية العامل العقلي الواحد، ولذلك رفض مفهوم الذكاء العام واستبدله بصور نوعية للذكاء وهي الذكاء المجرد. والذي يعني به القدرة على معالجة الألفاظ والرموز. والذكاء الميكانيكي يعني به القدرة على معالجة الأشياء والمواد العيانية. والذكاء الاجتماعي ويعني به القدرة على التعامل مع الآخرين. ويرى ثورندايك ان مزج الأنواع الثلاثة من الذكاء في الاختبار يؤدي إلى الارتباك والخلط.

أما طومسون فقد وجه انتقاده إلى نظرية العاملين ويرى طومسون أن أي نشاط عقلي للفرد يعتمد على عينة محدودة أو بمجموعة من هذه العناصر وينتج الارتباط من تداخل عينات مختلفة من هذه العناصر وبهذه الطريقة تنتج الأنواع المختلفة من العوامل التي تتدرج من العوامل النوعية إلى الطائفية المتباينة المدى إلى العامل الزائد الاتساع. ويرى طومسون أن التحسن الناتج في أي نشاط عقلي نتيجة للتدريب لا

يرجع إلى التحسن في القدرات الأولية المتضمنة، بل في استخدام وسيلة أفضل أو أكثر اقتصادا في انتقاء هذه القدرات.

أن طومسون لا ينكر وجود الارتباطات التي تدل على العامل العام، وإنما ينكر باعتبارها النوع الوحيد الموجود، وان سبب ظهورها في بحوث سبيرمان هو أن الاختبارات المستخدمة كانت قليلة العدد وتجمعها صفات مشتركة، فإذا زاد عدد هذه الاختبارات وتباينت خصائصها أمكن ظهور أنواع أخرى من التجمعات التي تربط بين بعض الاختبارات دون الاختبارات الأخرى. إن هذا النوع الأخير من الارتباطات يفسر على أساس وجود العوامل الطائفية وهي الناحية التي اهتم بها طومسون. ظلت نظرية العينات مجرد تصور نظري لا يوجد ما يدعمه من النتائج التجريبية رغم ما بذله طومسون لعمها بالأسس الرياضية والإحصائية على أساس افتراض معاملات الارتباط الموجبة بين الاختبارات العقلية.

أما لويس ثيرستون Thurstone فقد اقترح أن هناك ثمان من القدرات الأولية تقريبا والتي تعتمد نسبيا بعضها على البعض الآخر وكانت قدرات ثيرستون المقترحة هي:

- عامل الاستنباط.
- المنطق الاستقرائي
- عامل العلاقات اللفظية
- عامل طلاقة الكلمات
- عامل الاستدلال
- العامل العددي
- عامل السهولة في التصور المكاني والبصري
- عامل السرعة الإدراكية

وبعد أن تؤخذ القدرات العقلية الأولية بالاعتبار فإن مسألة العامل العام كانت غير مهمة نسبياً. وتوصل ثيرستون إلى نظرية جديدة عرفت بـ "نظرية العوامل المتعددة".

توصل ثيرستون في أبحاثه إلى ما يلي:

1- إن اختبارات الذكاء لا تقيس قدرة عامة واحدة بل كثيرا من القدرات أو العوامل الأولية كالقدرة العددية والقدرة اللفظية وغيرها.

2- إن هذه القدرات الأولية مستقل بعضها عن بعض استقلالاً نسبياً لا مطلقاً.

3- تتظافر هذه القدرات بعضها مع البعض في الإنتاج العقلي خاصة المعقد، فيندر أن يقوم الشخص بإنتاج معقد لا تؤثر فيه هذه القدرات جميعاً بنسب متفاوتة.

أما سيرل بيرت C. Burt فقد وضع في المستوى الأدنى عمليات الإحساس البسيط والنشاط الحركي البسيط، بعدها الإدراك والحركات التآزرية والتي تتضح في تجارب إدراك الأشكال والأنماط وفي الإرجاع الحركية، أما المستوى الثالث فهو المستوى الارتباطي حيث يضع بيرت عوامل الذاكرة والعادات التي تم اكتسابها وتكوينها، أما المستوى الرابع ففيه عمليات العلاقات وتنقسم إلى الثنائية العظمى في الفهم من ناحية والاستخدام من ناحية أخرى، أما الذكاء العام أو الإمكانية التكاملية للعقل فيظهر في كل مستوى من المستويات الأربع، إلا أن ظهوره يختلف في كل مستوى من حيث الدرجة والنوع. وأساس هذه الفكرة أن العقل لا يعمل على مستوى واحد من التعقيد.

أما كلفورد J. P. Gilord فيعتبر الرائد الأول لمفهوم "القدرات المضاعفة" Multiple abilities فقد ذهب - كلفورد - إلى أن العمليات العقلية يمكن أن تجزأ إلى ثلاثة مجموعات مستقلة.

أولهما: أن السلوك العقلي يتضمن بعض المحتوى (Content) فنحن نفكر في شيء: وان ذلك الشيء الذي يعرف محتوى السلوك العقلي عموما نحن نفكر في:

- المحتويات الرقمية Figural Contents
- المحتويات الرمزية Symbolic Contents
- محتويات تغير المعاني Semantic Contents
- المحتويات السلوكية Behavioral Contents

ثانيهما: أن تفكيرنا يتضمن الإجراءات Operations مثلا: أن الغرض من سلوكنا العقلي يمكن أن يتضمن:

- التقويم Evaluation
- الإنتاج المتقارب Convergent production (كتحديد الجواب الصحيح مثلا)
- الإنتاج المتباعد divergent production (مثل العصف الدماغي أو أن ندع العقل المبدع لأحدنا أن يسلك طريقه الخاص)
- الذاكرة memory
- أو الإدراك Cognition (مثل ما يعرفه الشخص أو ما يستطيع اكتشافه بسهولة)

ثالثهما: هناك منتجات products تنتج من عملياتنا العقلية أن هذه المنتجات يمكن أن تكون

- وحدات عقلية Mental Units
- مجا ميع Classes
- علاقات Relations
- أنظمة Systems
- تحويلات Transformations
- تطبيقات implications

من هنا ومن وجهة نظر كلفورد فإن كل مهمة عقلية تتعامل على الأقل مع واحد من المحتويات الأربعة، والعمليات الخمسة، والمنتجات الستة، ومن ثم فإن هناك على الأقل (120) تركيبة مختلفة ممكنة أو أنواع (أنماط) مستقلة من السلوك العقلي 6x5x4 =120.

أما فكرة المواءمة لتنظيم القدرات العقلية فقد وصفت من قبل فيليب فيرنون Phillip E. Vernon والتي يوضحها الشكل التالي: فقد تخيل فيرنون أن القدرات العقلية تنتظم تسلسليا حيث تقع القدرات الذهنية العامة في القمة ثم تنحدر منها مهارات أكثر نوعية.

مثلا: يتضمن المستوى الثاني القدرات اللفظية - التعليمية والتي هي بدورها تلحق بها مهارات أكثر تخصصاً من:

- الطلاقة اللفظية Verbal Fluency
- القدرات الإبداعية Creative abilities
- القراءة Reading
- التهجي Spelling
- المهارات اللغوية Linguistic Skills
- المهارات الكتابية Writing skills

التنظيم الثاني في هذه السلسلة أكثر عملية (من الأول) ويتضمن:

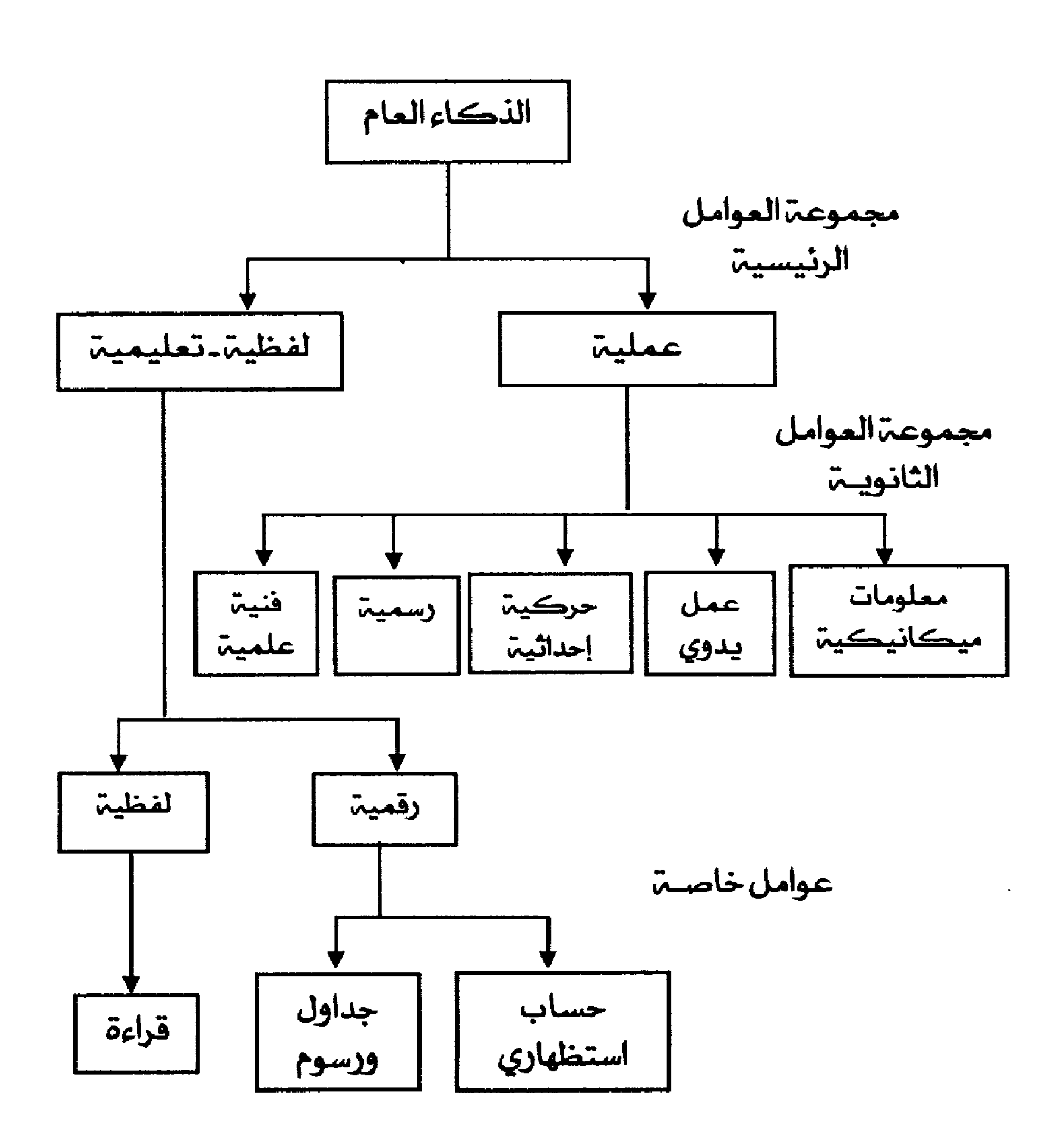
- القدرات الحركية Psychomotor abilities
- المهارات الفيزيائية (المادية) Physical skills
- المعلومات الميكانيكية Mechanical information
- القدرات الفراغية (المكانية) spatial abilities

... الخ وأخيراً تنحدر المهارات الحسابية Mathematical skills من الذكاء العام ومن القدرات اللفظية -التعليمية والقدرات العلمية. وبمدلول فيرنون فإن النجاح المدرسي غالبا ما يتألف مع الذكاء العام، إضافة إلى بعض المساهمة من المهارات اللفظية -التعليمية والقدرات الأكثر نوعية: القراءة، التهجي، الحساب.

من الناحية الأخرى فإن النجاح "كمهندس" ربما يتضمن الذكاء العام ومساهمة كل من القدرات اللفظية - التعليمية والقدرات العملية إضافة إلى مساهمات أخرى من القدرات الحسابية، العملية، الميكانيكية، والمكانية وإذا ما نظر من خلال وجهة نظر فيرنون فإن اختبارات ستانفورد بينيه وويكسلر تنزع إلى تمثيل الذكاء العام إضافة إلى تمثيلها لمساحات المهارات اللفظية التعليمية، الحسابية، والفراغية. إن اختبار ويكسلر يمتلك فقرات اختباريه أكثر تتعامل مع الجانب التطبيقي (العلمي)

لنموذج فيرنون، غير إنه أي منهما (أي اختباري بينيه وويكسلر) لا يظهر إنه يستطيع إخبارنا كثيرا من القدرات الحركية أو الإبداعية على سبيل المثال.

رسم توضيحي جزئي لمنظومة فيرنون للقدرات الإنسانية



ثبات نسب الذكاء

شعر كثير من الناس أن الذكاء الحقيقي يجب أن يكون ثابتا نسبيا عبر حياة الفرد فإذا ما أحرز الأفراد ذكاء عاما ممتازا إذ ذاك ربما يجب أن يحرزوا ذات الذكاء خلال حياتهم... (فنحن لسنا لامعين في عام وبليدين في عام آخر) لذا وإذا ما عكست اختبارات نسب الذكاء "الذكاء الحقيقي" بأمانة فإن درجة نسبة ذكاء الطفل يجب أن تكون متماثلة تقريبا عندما ننتقل به من تقدير لآخر، ولحد ما فإن هذا صحيح.

مثلا: إذا ما طبق اختبار نسبة الذكاء على مجموعة أطفال في سن العاشرة من العمر ثم أعطوا بعد أسبوعين ذات الاختبار أو نسخة مقاربة ذات أسئلة مختلفة فإن درجات هؤلاء الأطفال في كل من هذين الاختبارين ستكون متشابهة جدا، وعادة ما يكون الفرق من (6–8) درجات مابين الاختبارين. غير أن من الندرة أن يسجل احدهم (118) في أسبوع و(103) في الأسبوع التالي. وبالفعل فإن الكثير من الأفراد الذين اخذوا اختبار نسب الذكاء لعدة مرات خلال أعوامهم لم تتغير درجاتهم كثيرا جدا من سنة إلى أخرى وبالتقريب فإن ثلث المجتمع لا يظهر تغيرا يزيد عن درجات (أوطأ أو أعلى) من قيمة المعدل. وفي الحقيقة فإن الانجاز النسبي في اختبارات الذكاء ربما يكون واحد من الخصائص النفسية الأكثر استقرارا في كل ما قيس منها. ولكن عندما نقرا بأن ثلث المجتمع (خاصة أولئك الذين لا يسجلون عاليا في اختبار نسب الذكاء) ربما سيسجلون ذات الدرجة عاما بعد عام فإن الجزء الباقي من المجتمع ربما سيظهر تحولات واضحة في هذا الانجاز النسبي.

مثلا: في الدراسات التي أجريت حول تغير نسبة الذكاء (في معهد بحوث فيلس في مدينة يلوسبرنج في ولاية اوهايو في الولايات المتحدة الأمريكية، عبر شرائح من أطفال الطبقتين العليا والوسطى التي تعيش في تلك المنطقة أشاروا إلى أن ثلثهم قد تتذبذب درجات نسب ذكائهم بحدود (30) درجة مابين أعلى وأوطأ درجتين يسجلونها خلال طفولتهم، وان واحدا من كل سبعة أطفال يظهرون كذلك تذبذبا

بحدود (40) أو أكثر من درجات نسبة الذكاء. وقد عرف بعض الأطفال - النادرين- الذين تغيرت درجات نسب ذكائهم لغاية (74) درجة مما يستطيع -هذا التغير- أن يرفعهم من أوطأ 2٪ إلى أعلى 2٪ من أفراد مجتمعهم. غير أن مثل هذه التغيرات في نسب الذكاء لا يظهر أنها تغيرات عشوائية بسيطة لقيمة مفردة تسجل عاليا في اختبار ثم واطئا في الثاني ثم عاليا في الثالث وهكذا... بل أن ما يبدو هو أن الأطفال يظهرون أما تدرج تقدمي أو تأخري في درجات نسب ذكائهم وعلى امتداد أو مسافة بضع سنوات وان هناك تغيرات في اتجاه هذا التدرج النسبي (كأن ينتقل من الانحدار إلى الارتفاع في خط التغير أو العكس) مابين السنتين السادسة والعاشرة من العمر.

خلاصة القول

هو أن الكثير من الأفراد لا يتغير انجازهم النسبي في اختبار نسبة الذكاء كثيرا جدا وأن آخرين منهم يظهرون تحولات واضحة في وضعهم النسبي. إن هذه التحولات في مواقع الأفراد النسبية تبين لنا أن من الخطأ أن يفكر أو يعتقد احدنا أن الفرد قد قفل عليه (أو احتجز) بمستوى معين من الانجاز النسبي طيلة حياته حتى عندما يستخدم اختبار نسب الذكاء لغرض التنبؤ بالنجاح المدرسي فإنه يظهر أن من الحماقة إعطاء اختبار ذكاء واحد فقط واعتماد نتائجه لتحديد الخبرة والفرصة التعليمية للطفل لما يتبقى من حياته. وهذا صحيح بالذات عندما نتذكر أن التقلبات الواضحة في الانجاز النسبي تنزع إلى الظهور مابين السنة السادسة والعاشرة.

التحيز الثقافي Cultural Bias

في امتحان قريب لفقرات اختباريه مأخوذة من اختبار نسبة الذكاء أشارت النتائج إلى أنها - أي الفقرات- نموذجية أو ملائمة بدرجة كبيرة إلى الأمريكيين من أفراد الطبقة المتوسطة أكثر بكثير من ملائمتها لأفراد الأقليات الثقافية التي تعيش في الولايات المتحدة الأمريكية كالسود الفقراء، البورتيريكين، الهنود الحمر... الخ.

حيث يظهر الطفل الأبيض في أعلى الطبقة الوسطى إنه أكثر مشاهدة للمضلة مثلا، ومن ثم فهو غير قادر على تعريف هذه الكلمة أكثر مما يستطيعه الطفل الذي يعيش مثلا في إحياء اليهود (Ghetto) في المقابل فليس هناك في الاختبار من مصطلحات معينة مثل الفأر وأمعاء الخنزير Caitlin والتي هي في غاية الشيوع في خبرة الطفل الفقير. وقد أظهرت الدراسة أيضاً أن وسائل الاتصال غير اللفظية Nonverbal Communication ذات أهمية اكبر عند اليهود مما هي في مجتمع الطبقة فوق المتوسطة وان الاختبار يركز بالتأكيد على العمليات اللفظية المجردة Abstract verbal processes في مقابل غير اللفظية منها وبالتالي فإن ما يجادل به الكثير هنا هو أن في الاختبار تمييز وتحيز واضحين ضد ثقافات أقلية معينة في الولايات المتحدة الأمريكية. إن السؤال المطروح هنا هو فيما إذا كان الاختبار مؤلما في الحقيقة للأفراد الذين لم يترعرعوا في جوهر التيار أو الفكر الأمريكي وبدون شك فهي تفعل ذلك لحد معين، ولكن التحيز - موضوع البحث- يمكن أن يكون أكثر تعقيدا مما يدعيه البعض.

مثلا: إذا ما نظرنا فقط إلى الفقرات الاختيارية التي تظهر أنها أكثر تحيزا لصالح أطفال الطبقة الوسطى فإننا نجد أن انجاز الأطفال الفقراء في هذه الفقرات - المتحيزة - اسوأ بكثير من إنجازهم في فقرات غير مشحونة بوضوح بصيغة ثقافية معينة (أي اقل تحيزا من الأولى) ولنفترض إزاحة الفقرات الاختيارية التي يعتقد أنها متحيزة للطبقة الوسطى في الاختبار، فإن أطفال الطبقة الوسطى يستمرون بتفوقهم على الأطفال الفقراء في الفقرات التي يفترض أنها غير متحيزة ثقافيا. وأخيرا إذا ما بنيت اختبارات خاصة نحاول أن تكون غير متحيزة ثقافيا. نسبيا فإن الأطفال الطبقة الوسطى والأطفال الفقراء يختلفون تقريبا بذات المقدار الذي يختلفون فيه باختبارات نسب الذكاء التي يعتقد أنها متحيزة ثقافيا.

والآن: هل يعني هذا أن اختبارات نسب الذكاء تقيس شيئا لا يتأثر بالخبرة؟

الجواب: لا، ليس الضرورة.

إذ ربما أن الأطفال الفقراء على وضع غير مؤات أو ملائم فيما يتعلق باختبارات نسب الذكاء المعيارية ليس لأنهم تعوزهم المعلومات المعينة والخبرة بل لأنهم لم يتدبروا بطريقة تنمي أو تطور عندهم المهارات الأساسية العامة الوثيقة الصلة بأنواع مختلفة كثيرة من المهمات والأسئلة.

ولنفترض أن هناك تحيزا ثقافيا في اختبارات نسب الذكاء: هل هذا يعني أننا يجب أن لا نتقدم خطوة لإبداع اختبارات تكون حقيقية غير متحيزة ثقافيا؟ يظهر أن الجواب يعتمد على ما إذا أراد احدنا استخدام اختبار نسب الذكاء كمقياس للذكاء الحقيقي- والذي يعتقد إنه غير متحيز ثقافيا أو إذا أريد استخدامه للتنبؤ بمجموعة أنماط سلوكية معينة كالنجاح المدرسي أو الوضع المهني (الوظيفي).

ما يعتقده بعض الناس هو أن الذكاء يجب أن لا يعكس ما تعلمه الطفل خلال نموه إنما يجب أن يقيس، بطريقة أو بأخرى، شيئا مستقلا عن خبرة الطفل وعندما يقال أن الذكاء لا يعتمد على خبرة الطفل فإن بذلك يقابل اقتراحا بأن يكون وزن الأفراد مستقلا عما يأكلونه خلال حياتهم.

جادل بعض العلماء في هذا فقالوا أن اختبار الذكاء غير المتحيز ثقافيا أو غير المتحيز للخبرة إنما يشبه لحد كبير مقياس وزن غير متحيز للطعام من الناحية الأخرى إذا ما أستخدم أحدنا اختبار نسبة الذكاء ليس غيره للتنبؤ بالانجاز الدراسي في المدرسة فإنه يتعين عليه أن يتوقع أن الاختبار سيكون على ذات التحيز بالنسبة إلى النظام التعليمي فإذا ما تحيز النظام التعليمي نحو نمو قيم وخبرات الطبقة الوسطى فإن من المعقول فقط حينذاك هو بناء اختبار للتنبؤ بالنجاح في هذا النظام يكون على ذات الطريقة من التحيز. هناك شك قليل في أن النجاح في التعليم المعاصر في أمريكا وربما في معظم الأقطار الغربية إنما يعتمد على العمليات اللفظية، المنطق المجرد، الذاكرة، التفكير الكمي، وبالضبط هذه هي المهارات التي تمثلها اختبارات نسب الذكاء المعيارية بكل قوة. فإذا ما أخذ التحيز الثقافي بطريقة أو بأخرى خارج النظام التعليمي الأمريكي، فإن من الممكن بعد ذلك أن لا تتنبأ

اختبارات نسب الذكاء بالنجاح المدرسي، وأنها حينذاك ستتغير بدون أدنى شك. إن الافتراض القائل بتحيز اختبارات نسب الذكاء ثقافيا لصالح أطفال الطبقة الوسطى ربما يكون صحيحا ولكنه - أي هذا التحيز- ليس إلى الدرجة التي يعتقد بها البعض في الواقع وإذا ما كان الاختبار مؤكدا بتحيزه الثقافي فإن أحدنا قد يدهش بقوة عندما يشاهد أطفالا من بيئات لا يتوقع منها النجاح يسجلون درجات نسب ذكاء عالية، ومع ذلك فإن كثيرين من هؤلاء الأطفال يحصلون على ذلك. وعلى كل حال فإن التحيز الثقافي لا يكون مشكلة إلا عندما نرغب من الاختبار أن يقيس شيئاً غير متحيز لخبرة يسمى "الذكاء" فإذا ما تقبل أحدنا اختبار نسب الذكاء، ابتداءً، كأداة عملية تستطيع التنبؤ بنجاح الفرد في فعاليات متحيزة ثقافيا، فأننا قد نتوقع حينئذ أن الاختبار قد يكون متحيزا بقدر ما يتحيز به المعيار.

الوراثة:

يركز هذا الجزء على كيف يمكن أن نعرف إذا ما تألف السلوك الذهني مع ما نرثه، ما هو الدليل الرقمي لمتوسطات الوراثة؟ وماذا يمكن أن نستخلص من السلوك الذهني إذا ما كان الموروث قويا أو ضعيفا.

فعندما نقول أن الذكاء يورث بقوة فإن هذه العبارة إنما تعكس الموروث من "الذكاء الحقيقي" إلا عندما يرغب الفرد أن يتقبل درجات اختبار نسب الذكاء كمقياس أمين "للذكاء الحقيقي" ومادام الناس يتقبلون هذا التفسير لدرجات نسب الذكاء والآخرون لا يتقبلوه فإن ما سيلي هو توضيح لهذه المواضيع.

أ- متى تتأثر الخصائص الإنسانية بالوراثة؟

إن الدراسات الحديثة لعلم الوراثة تتضمن على الأقل عنوانين رئيسيين:

أولهما: دراسة الانتقال الوراثي Genetic Transmission العمليات الكيمياحياتية Biochemical والفسيولوجية Physiological المتضمنة مرور أو انتقال المواد والسمات

الوراثية من عضو لآخر، وهو الجانب من الوراثة الذي يعتمد على المفاهيم الأساسية لانشطار الخلية Cell division، الجينات "ناقلة الصفات الوراثية" Genes الكروموسومات (مولدات الجينات) Chromosomes التغيرات Mutation، الخصائص المهيمنة مقابل الخصائص المنحسرة (المتغلبة والمتنحية) وهلم جرا.

العنوان الثاني هو علم الوراثة السكاني Population genetics والذي يتضمن التطبيقات العملية لمفاهيم الانتقال الوراثي في دراسة الأنماط والتكرارات للخصائص الموروثة في المجتمعات الإنسانية. وبينما يمكن للعلماء أن يجروا التجارب في انتقال السمات من عضو دوني لآخر فإنه من الوضوح بمكان عدم إمكانيتهم التجريب في مجال التزواج في المجتمعات الإنسانية. لذا فهناك حاجة لإجراءات خاصة، في حالة اكتشاف خاصية يختلف فيها الناس من فرد لآخر، لتحديد إن كان الاختلاف بينهما يعود للفروق فيما ورثوه أو بسبب الفروق في خبراتهم. وفي هذه: كيف يمكننا التحديد؟

ب ـ التشابه في نسب ذكاء ذوي القربى

يمكن لأحدنا أن يتحقق عن التعاون أو المساهمة الممكنة للوراثة في سمة ما عن طريق ملاحظة ما إذا كان الأفراد الأقارب أكثر تشابها بعضهم للبعض الآخر في تلك السمة نسبة إلى الأفراد الغرباء. فإذا تأثرت نسبة الذكاء بالصفات الموروثة فإن الأفراد المتقاربين وراثيا يجب أن يكونوا أكثر تشابها في نسب الذكاء عن الأفراد الغرباء والأكثر صلة للقربى يجب أن يكونوا الأكثر تشابها في نسب الذكاء.

على هذا النحو، فإن التوائم المتماثلة والذين يملكون ذات المجموعة من الجينات (الناقلة للصفات الوراثية) يجب أن يكونوا أكثر تشابها في نسب الذكاء من التوائم غير المتماثلة والذين هم يشتركون /في المتوسط/ بنصف مقدار الجينات التي يحملونها. والجدول أدناه يوضح درجة العلاقة بين درجات نسب الذكاء لأناس يملكون مقادير مختلفة من التشابه الوراثي، حيث يقدم الجدول الارتباط الوسيط

Median Correlation (على سبيل المثال الدرجة النموذجية للعلاقة) لكل نوع من أنواع صلات القربى بين الأفراد.

جدول يوضح ارتباطات نسب الذكاء لصلات قربى مختلفة

صلة القربى	الارتباط الوسيط
أطفال غرباء ترعرعوا منفصلين	-0,01
أطفال غرباء ترعرعوا سوية	0,24
الآباء وأبنائهم	0,50
أشقاء ترعرعوا منفصلين	0,47
أشقاء ترعرعوا سوية	0,55
توائم غير متماثلة (بجنس مختلف)	0,49
توائم غير متماثلة (ذات الجنس)	0,56
توائم متماثلة ترعرعوا منفصلين	0,75
توائم متماثلة ترعرعوا سوية	0,87

ولفهم هذا الجدول فإنه قد يحتاج إلى فهم شيء عن معامل الارتباط فالارتباط تعبير رقمي لدرجة التوافق بين مجموعتين من القياسات قدرت لذات الأفراد أو الأفراد ذوي علاقة.

فالارتباط (0,00) يمثل عدم وجود علاقة خطية مثلا:

في الجدول أعلاه فإن الارتباط مابين نسب الذكاء للأطفال الغرباء الذين ترعرعوا منفصلين (بعيدا عن بعضهم) هو (-0.01) أساسا مساويا إلى الصفر. وهذا يعني أن معرفة درجة نسبة الذكاء لطفل ما سوف تخبرنا أي شيء عما ستكون عليه درجة نسبة الذكاء لنظيره الزوجي (الفرد الثاني في الزوج المقارن) وهذا يعني أن لا علاقة هناك مابين نسبة الذكاء لفرد ما ونسبة ذكاء فرد آخر غريب عنه –أي

عن الفرد الأول - من الناحية الأخرى، لنفترض وجود ارتباط (1,00) حينئذ فإن التوافق النسبي مابين الأزواج المقارنة ستكون تامة وفي هذه الحالة فإن معرفتك لطفل حصل على أعلى درجة (أو أوطأ درجة) نسبة ذكاء في مجموعته سوف تشير إلى أن نظيره الزوج يحصل على أعلى (أو أوطأ) درجة نسبة ذكاء في مجموعته. ما يلاحظ هنا أن معامل الارتباط يعكس الوضع النسبي ضمن المجموعة فقط.

إن الارتباط التام (1,00) يخبرنا إنه إذا ما كان طفل ما نسبة ذكائه (136) هو الأعلى أو (أوطأ) في مجموعته فإن نظيره (في الزوج المقارن) سيكون هو الأعلى (أو إلا وطأ في نفس المجموعة) ولكن لا يحتمل أن يمتلك هذا النظير ذات نسبة ذكاء الأول تماما: أي 136.

في البحث النفسي نادرا ما يكون مساويا إلى الصفر (0,00) أو واحدا صحيحا (1,00) بشكل أدق منها - أي معاملات الارتباط- تقع في أماكن مختلفة مابين هذين الطرفين، والأكبر في الحجم الرقمي المجرد للارتباط هو الأكثر درجة للتوافق مابين الوضع النسبي لأحد أفراد الزوج - والوضع النسبي لنظيره الزوجي ضمن مجموعاتهم. ولنرجع الآن بالاعتبار إلى الجدول أعلاه أن القيم المعروضة في هذا الجدول يجب أن تؤخذ بصورة تقريبية كذلك حيث يمكن ملاحظة إنه كلما زادت درجة القرابة بين أفراد أزواج المقارنة فإن الارتباط بين نسب ذكائهم تزداد كذلك..

مثلا: الارتباط بين التوائم المتماثلة التي ترعرعت سوية هو (0,87) بينما الارتباط بين التوائم غير المتماثلة (0,56) ومادامت التوائم تشترك بكل جيناتها والتوائم غير المتماثلة تشترك بنصف هذا المقدار (من الجينات) فإن هذا الفرق يمكن أن يتوقع استنادا أو على أساس درجة تشابههم الوراثي، مع ملاحظة أن الأشقاء (الأخوة والأخوات) الذين يترعرعون سوية، والتوائم غير المتشابهة والآباء وأبنائهم كلهم يقع ارتباطهم مابين (0,49) و(0,56) والتي تعتبر هذه المجموعات المقارنة- نسبة إلى الحقائق الوراثية أن أفراد كل من هؤلاء الأزواج يشتركون بالمتوسط بنصف عدد جيناتهم. من الناحية الأخرى، فإن الأطفال الغرباء الذين

ترعرعوا سوية يمتلكون ارتباطا قدره (0,24) ومادام هؤلاء الأطفال يفترض أو يسلم بعدم مشاركتهم لأي جينات، فمن وجهة النظر الوراثية فإن ارتباطهم يجب أن يكون مساويا إلى الصفر تماما والمفروض هنا هو أن الـ (0,24) إنما تعكس حقيقة أن أفراد هذه الأزواج – الغريبين عن بعضهم - قد ترعرعوا في ذات البيئة كما يمكن أيضاً ملاحظة أن الارتباط مابين الأشقاء الذين ترعرعوا سوية أعلى من الارتباط للأشقاء الذين ترعرعوا منفصلين (0,55 و 0,47 على التوالي) وهي – الارتباطات - أعلى كذلك عند التوائم المتشابهة التي ترعرعت سوية مما هي عند التوائم المتشابهة التي ترعرعت بعيدا عن بعضها (0,87 و 0,75 على التوالي) مما يمكن أن تؤدي بنا – هذه الملاحظات – إلى التوافق مع فكرة تأثير البيئة على الذكاء.

ج- التشابه الوراثي والبيئي

في تفحص غير ناقد للارتباطات التي يبينها الجدول يظهر ما يشير إلى خلاصة مؤداها أن التشابه في إنجاز اختبار نسب الذكاء يتغير مباشرة (طردا) مع التشابه الوراثي خلال أفراد الأزواج و من الناحية الأخرى فإن بعض العلماء جادلوا بأن هذه الارتباطات غامضة ولا ترى بالضرورة أسساً وراثية لنسب الذكاء حيث أشاروا – هؤلاء- إلى إنه كلما زاد التشابه الوراثي زاد التشابه البيئي ومن ثم فإن التشابه الملاحظ ضمن الأزواج ربما يعود حقيقة إلى التشابه في البيئة – لهذه الأزواج- وليس إلى الجينات المشتركة عندهم ولنأخذ شاهدين نوعيين على ذلك

أن اغلب التقدير – القياس- المباشر للمساهمة الوراثية والبيئية للانجاز في اختبار نسب الذكاء عادة ما يتضمن مقارنة للتشابه في نسب الذكاء خلال أزواج من التوائم المتماثلة مقابل التوائم غير المتماثلة. ومادامت التوائم المتماثلة تشترك بكل جيناتها في الوقت الذي لا تشترك فيه التوائم غير المتماثلة في المتوسط إلا بنصف هذه الجينات فقط، فإنه ليس مدهشا – من وجهة النظر الوراثية – أن يكون ارتباط نسب الذكاء للتوائم المتماثلة (0,87) مقابل (0,56) للتوائم غير المتماثلة من ذات الجنس (ذكور فقط أو إناث فقط).

من الناحية الأخرى فإن الناقدين جادلوا-بالنسبة للشاهد النوعي أعلاه - بأن الآباء يعاملون توائمهم المتماثلة بتشابه أكثر (بين التوأم والأخر) مما يعاملون به توائمهم غير المتماثلة.

ومن ثم فإن البيئة أكثر تشابها لأزواج التوائم المتماثلة مما هي لغير المتماثلة، وان التشابه الأكثر في نسب الذكاء للتوائم ربما جاءت من التشابه البيئي الكبير لهم وليس من خلال البنية الجينية المشتركة لهم. المفروض أن مثل هذه الاحتمالية يمكن اختبارها متى ما قورنت التوائم المتماثلة التي ترعرعت سوية مع التوائم المتماثلة المترعرعة في بيئات مختلفة. فإذا ما كانت الجينات هي العنصر الرئيسي في تشابه نسب ذكاء أفراد الأزواج فإن ارتباطات التوائم المتماثلة المترعرعة في بيئات مختلفة يجب أن تكون حينئذ مقاربة للتوائم المترعرعة في بيئة واحدة. على عكس ذلك فإن هذه الارتباطات يجب أن تكون مختلفة من ارتباط لآخر لمدى الدور أو الوظيفة التي يمكن أن تلعبها البيئة. الارتباطات الفعلية -كما يريها الجدول هي (0,87) للتوائم المترعرعة سوية (0,75) للتوائم المترعرعة منفصلة ومع أن هذه الارتباطات ليست بذات القيمة (مما يظهر أن للبيئة هنا بعض التأثير) إلا أنها كذلك ليست على درجة كبيرة من التباين.

الناقدين لهذا السبيل - من فهم أو فسر النتائج- يجادلون بأنه عندما تنفصل هذه التوائم المتشابهة عن بعضها (لتترعرع في بيئات مختلفة) فإنها ما توزعت قط تقريبا في بيئاتها البيتية بطريقة عشوائية. وفي الحقيقة فإن التوائم المتماثلة غالبا ما توضع في بيوت الأعمام أو العمات، الأخوال أو الخالات، أو الجدود.

وما نتوقعه فإن الجو الفكري العام يكون متشابها نسبيا في بيوت مختلف الأفراد من ذات العائلة، وبالتالي فإن التوائم المتماثلة المترعرعة منفصلة لا تعرض / في الحقيقة/ لبيئات مختلفة جدا - ربما ليس أكثر اختلافا مما لو أنها لم تنفصل عن بعضها.

كانت هناك في بعض الدراسات محاولات لقياس مدى اختلاف البيئات التي تعيشها التوائم المنفصلة، إلا أن المشكلة كانت في أن علماء النفس ليسوا غاية في التأكد من الجزء الفعلي من البيئة الذي يؤثر في نمو مهاراتهم الفكرية المقاسة باختبارات نسب الذكاء ومن ثم، فقد جادل بعض العلماء من إنه حتى لو كانت البيئات المنفصلة التي يترعرع بها التوائم مختلفة جدا من واحدة إلى أخرى في هذه القياسات فنحن لا نملك الضمان بأن الصفات المميزة للبيت المنعكسة في هذه التقديرات البيئية تساهم فعلا في انجاز اختبار نسب الذكاء.

يمكن أن يجري جدل مقارن حول التشابه البيئي عندما يقارن الأشقاء بالأطفال الغرباء شأنها شأن صلات القربى الأخرى. في الحقيقة فإن بعض العلماء قد تركت عندهم هذه الفوضى (أو الخلط) بين التشابه البيئي أو الوراثي انطباعا قويا جعلهم يقترحون إنه ليس هناك من دليل قط على أن الانجاز في اختبار نسب الذكاء إنما يعود إلى الوراثة بأية طريقة كانت وعندما تقع الحقيقة ربما في موقع ما سوى هذا الموقع المتطرف فإن ليس من المدهش كثيرا أن تثار هذه المناظرة.

د- مفهوم الموروث

بالإضافة إلى حقيقة أن التشابه الوراثي والبيئي ينزع إلى السير سوية ولا يمكن فصلهما عن بعضهما بسهولة ففهم الناس للشاهد أو الدليل على المساهمة الوراثية في درجات اختبار الذكاء قد تعقدت بزيادة أكثر بسبب الطبيعة التقنية التي يعبر بها العلماء عن المساهمة الوراثية في تلك السمة مثلا: بعد مراجعة التأليف العلمي في الوراثة ونسب الذكاء فإن آرثر جينس Arthur Jensen قد توصل إلى أن الموروث من نسب الذكاء يقارب (0,80) فما هو الموروث وكيف يمكن تفسيره؟

ما هو الموروث:

الموروث - في ابسط صيغة له - مفهوم إحصائي يعكس لنا نسبة التغير في سمة (أو صفة) تتآلف مع الفروق في التركيب الوراثي لأفراد في مجموعة.

مثلا: افترض إنه تم إعطاء اختبار نسب الذكاء إلى (100) فرد. هؤلاء الأفراد سوف لا يسجلون كلهم ذات القيمة في هذا الاختبار (الفروق بينهم - الأفراد - في هذه المجموعة تسمى "التغاير" Variability) لاشك أننا نعرف أن أفراد هذه المجموعة يمتلكون تركيبات وراثية مختلفة شانها شان تباين خبراتهم الحياتية. وما نرغب فيه هو معرفة كم من هذا التغير في درجات نسب ذكائهم تتألف مع الفروق في بنيتهم الوراثي. إن هذه النسبة المئوية هي الموروث من نسب الذكاء لهذه المجموعة من الأفراد.

ما الذي نعنيه بالموروث؟

الموروث تطبيق للمجموعات وليس للأفراد أن أول شيء نحتاجه للتمييز هنا هو أن قيمة الموروث تطبيق أو استخدام للفروق في نسب الذكاء لمجموعة من الأفراد وليس تطبيق لنسبة ذكاء أي عنصر مفرد. فإذا ما كان الموروث من نسب الذكاء (0,80) فإن هذا يعني أن (80٪) من الفروق بين الأفراد تتألف من الفروق في الخصائص الموروثة لهؤلاء الأفراد. ومن المفترض هنا أن الـ (20٪) المتبقية تتألف مع خبرات الحياة المختلفة. ومع ذلك وإذا ما كان الموروث (0,80) فإن هذا لا يعني إذا ما كان فرد ما يمتلك نسبة ذكاء (100) فإن (80٪) أو (80) درجة منها ترجع إلى الوراثة وان (20) نقطة منها قد ساهمت بها البيئة. فالموروث يرجع إلى الفروق بين عناصر مجموعة من الناس وليس إلى القيمة المجردة لدرجة الفرد.

أن الموروث يعتمد على العينة المتوفرة من الأفراد وعندما يحدد الموروث لسمة معطاة، فإنه يقدر على عينة من الأفراد. قد يأتي هؤلاء لأفراد من بيئات مختلفة أو بيئات جدا متشابهة، وربما يكون تركيبهم الوراثي مختلفا جدا أو متشابها جدا من فرد لآخر. إن الحجم الرقمي للموروث إنما يعتمد على تباين البيئات والخلفيات الوراثية الممثلة من قبل الأفراد في تلك المجموعة المعينة.

خذ المثالين المتطرفين التاليين بالاعتبار:

- افترض أن احد استطاع أن يحصل على مجموعة من الأفراد كلهم يمتلكون ذات البنية الوراثية ولكنهم قادمون من بيئات فكرية متنوعة. فإذا ما أعطى هؤلاء لأفراد اختبارا لنسب الذكاء فإنهم سوف لن يسجلوا جميعا ذات الدرجة. ومادام جميعهم يمتلكون ذات البنية الوراثية فإن أياً من الفروق في نسب الذكاء بين الأفراد لن يكون متآلفا مع بنيتهم الوراثية وبالتالي، فإن الموروث من نسب الذكاء في هذه العينة سيكون مساويا إلى الصفر.
- في اليد الأخرى افترض إمكانية الحصول على عينة من الأفراد ممن يختلفون بشكل كبير في تركيبهم الوراثي ولكنهم بطريقة أو بأخرى ترعرعوا بظروف متماثلة (متطابقة). هؤلاء الأفراد سيختلفون أيضاً في نسب ذكائهم ولكن كل الفروق في قيم درجاتهم (نسب ذكائهم) في هذه المرة ستكون متآلفة مع الفروق في تراثهم الوراثي لذا فإن الموروث من نسب الذكاء في هذه العينة سيكون مساويا إلى (1,00).

ما توضحه هاتان الحالتان المتطرفتان هو أن الدليل الرقمي للموروث يعتمد على المدى الذي يتغير به أفراد العينة في ظروفهم الوراثية والبيئية. وهذا شيء من الأهمية بمكان لان كثير من العينات التي استخدمت لتحديد الموروث من نسب الذكاء تنزع في تكوينها إلى البيض من الطبقة المتوسطة والذين يتشابهون أكثر مما في أصولهم الوراثية.

لقد افترض بعض العلماء أن الموروث من نسب الذكاء سيكون اقل اعتبارا إذا ما مثلت بيئات ذات تباين أكثر (كاشتمالها على حالات الفاقة والعوز).

لقد استخدمت الكثير من الدراسات الخاصة بالموروث من الذكاء عينات بريطانية، والذي يظهر أن الموروث من نسب الذكاء في انكلترا أعلى مما هو في الولايات المتحدة الأمريكية. وقد توقع بعضهم أن هذا - الاختلاف - بسبب تنوع وتباين الظروف البيئية في الولايات المتحدة أكثر مما هو جار في انكلترا. ولا يختلف الموروث باختلاف العينات فقط إنما قد يتغير من سنة إلى سنة تالية لعينات مقارنة.

خذ حالة التدرن (السل) بالاعتبار (Tuberculosis). قبل سنوات مضت كان الموروث من السل الرئوي (TB) عاليا جدا وكان السبب هو أن عصيات السل الرئوي كانت من التفشي بمكان بحيث ما من فرد تقريبا إلا وكان على تماس بها. المحدد الرئيسي للإصابة بالسل الرئوي كانت الحساسية الكيمياحياتية الموروثة لهذا المكروب وبالتالي فإن الفرق مابين من أصيب بالسل الرئوي ومن لم يصب كان متآلفا بقوة مع الظروف الوراثية، وحيث كان الموروث منه عاليا.

أما ألان، من الناحية الأخرى فإن تقدم الوقاية الصحية العامة قد استأصل هذه العصيات تماما ماعدا البيئات الصحية البائسة ومن ثم فإن جزءاً قليلا من السكان فقط والى حد بعيد يعرض لهذا الجرثوم وبالتالي فإن الفرق مابين من أصيب بهذا المرض ومن لم يصب به يتآلف بصورة اكبر مع حياتهم البيئية مما هو مع الحساسية الكيمياحياتية الموروثة. هناك الكثير من الأفراد من حولنا من له سرعة التأثر بالجرثوم ولكنه لم يتماس معه أو يصيب بالسل الرئوي. والآن فإن الموروث من السل الرئوي قليل جدا هذه الأمثلة توضح أن الموروث من نسب الذكاء يعتمد على العينة الجاهزة خصوصا ذات البيئات والبنيات الوراثية المتباينة عند أفرادها.

الفرق بين الذكاء والاستعداد والقدرة

استخدم بعض العلماء مصطلحات الذكاء Intelligence والاستعداد Aptitude والقدرة Ability بالتبادل أحدهما محل الآخر في كثير من الدراسات، رغم ما بينها من فروق قليلة ودقيقة في تمييز معناها. ويبدو منذ السنوات الماضية أن علماء النفس يستخدمون الذكاء والقدرة والاستعداد في وصف وتمييز التكوين أو مجموعة التكوينات التي قد تكون مفيدة في المساعدة على التنبؤ بالأنواع المختلفة للسلوك، ولذا فإن الاختبارات المصممة لقياس هذه التكوينات تختلف بعض الشيء باختلاف التكوينات المراد قياسها.

ويرى كرونباخ أن اختبارات القدرة والذكاء والاستعداد تقع جميعها تحت اختبارات أقصى الأداء الممكن Maximum Performance والتي تهدف إلى معرفة مستوى أداء الفرد في أفضل ظروف ممكنة بحيث ان الفرد سيبذل أقصى جهد في الاختبار.

أن التمييز بين الذكاء والاستعداد لم يكن عموما واضحاً، بيد أن الذكاء بصورة عامة تكوين أوسع من الاستعداد، وقد ذكر كليري Cleary عام 1975 أن الذكاء عبارة عن مستودع عام لما يكتسب من خبرات التعلم والمعرفة والمهارات. كما أن الأهداف العامة لهذه الاختبارات تعتبر بطبيعتها عقلية في الفترة المعينة لإجراء الاختبار. في حين أن الاستعداد كمفهوم يشير للخصائص النفسية التي تعده قبلاً وتتنبأ بالاختلافات في التعلم المتأخر تحت حالات تعلمية خاصة كما أن معظم الكتاب يشعرون بأن مهمة إيجاد صدق البناء (التركيب) في اختبارات الذكاء أكثر تحدياً من إيجاده في اختبارات الاستعدادات حيث أن معظم اختبارات الذكاء لا تحصل على درجات عالية في صدق البناء لتميز الذكاء

إن الاستعداد هو الناحية التنبؤية للقدرة، فهو التجمع المتناسق للصفات والخواص التي تدل على استطاعة الفرد للقيام بعمل معين أو نمط محدود من أنماط السلوك، أو إنه القدرة الكامنة للسلوك، معنى ذلك أن الاستعداد سابق على القدرة وملازم لها، والقدرة ما هي إلا قدح للاستعداد الخاص بعوامل البيئة والنضج. فالفرد الذي عنده استعداد موسيقي معنى ذلك أن لديه قدرة خاصة إذا تعهدناها بالتدريب أمكنه التفوق فيها، فكأن الاستعداد قدرة موجودة لدى الفرد في حاجة إلى التنمية بالتدريب.

كما أن الاستعداد هو نتيجة تفاعل الوراثة والبيئة معاً، ويتضمن الشمول من حيث الذكاء والتحصيل والشخصية والميول وأي قدرات أو مهارات أخرى ولا يقتصر على القدرات الخاصة فقط.

كما أن اختبارات الذكاء أو اختبارات القدرات العقلية يمكن استخدامها في التعرف على القدرة في التعلم المدرسي ولهذا فإن الذكاء والاستعداد والقدرة أكثر تشابهاً فيما تهدف إليه من اختبارات التحصيل، إلا أن الكثير من علماء النفس في الوقت الحاضر يجدون بعض التشابه الكبير بين اختبارات الذكاء والتحصيل فيما تقيسه من مهارات وقدرات مدرسية.

وأشار Caroll 1974 إلى أن هناك بعض التمييز بين الاستعداد والتحصيل من حيث الخبرة السابقة التي يقيسها كل منهما، حيث يرتبط التحصيل بقياس مجموعة من الخبرات الموحدة والمقصودة نسبياً، أما الاستعداد فيرتبط بالآثار التراكمية لخبرات غير متقنة وغير مقصودة في الحياة اليومية نسبياً، كما ان اختبارات الاستعدادات تفيد في التنبؤ في الأداء اللاحق في حين تستخدم اختبارات التحصيل في تقويم آثار برنامج التعليم أو التدريب.

نماذج من اختبارات الذكاء

اختبار ستانفورد-بينيه التنقيح الثالث -صيغة 1960

Stanford-Binet Intelligence Scal By, Lewis M. Terman and Maud A. Merrill Third Revision Form L-M (1960)

السمات العامة

- يتمتع التنقيح الثالث بذات السمات والملامح للطبعات السابقة مع حذفها لبعض الفقرات التي أصبحت قديمة وغير مناسبة للوقت الحاضر.
- التنقيح الجديد في صيغة واحدة (L-M) (بعد أن كان في صيغتين لطبعة 1937. التنقيح الثاني (L&M) جمعت معظمها من الفقرات المميزة والجيدة للصيغتين التي احتواها التنقيح الثاني.
- ابتدأ التنقيح الثالث لويس تيرمان بنفسه وقاد التجارب الأولية العامة التي

ابتدأت لأجل هذا التنقيح للسنوات (1950-1954) ولكنه توفي (1956) قبل ظهور التنقيح الثالث بصورته النهائية في عام 1960 فأكملها بعده شريك عمله في جامعة ستانفورد الدكتور مودميرل.

نظرة تاريخية

(أ) الطبعة الأولى (1916)

- ظهرت الطبعة الأولى لمقياس ستانفورد بينيه في عام 1916 من قبل لويس تيرمان مستندة إلى مقياس بينيه وسيمون الفرنسي الأصل.
- ما حاوله المقياس الجديد هو وضع معايير للانجاز العقلي مناسبة للأطفال المولودين في الولايات المتحدة الأمريكية من عمر الثالثة إلى عمر السادسة عشرة.
- رتبت اختبارات هذا المقياس ودرجت بحسب صعوبتها للمستويات العمرية
- يحكم على القدرة العقلية للفرد المحددة بالانجاز في هذا المقاس من خلال المقارنة مع معايير الإنجاز للأطفال العاديين بمختلف الأعمار أن احد افتراضات بينيه الأساسية هو أن الشخص يعتبر عاديا إذا استطاع أن يقوم بما يقوم به الأفراد الآخرين من ذات عمره وهو متخلف أن تدنى عن المتوسط ومتقدم أن تقدم عليه.
- في هذا المقياس زاد عدد الفقرات الاختبارية إلى (90) بعد أن كان (54) في اختبار بينيه الأصلي مع الاحتفاظ بذات الأنماط الاختبارية التي يتألف منها الاختبار الأصلي.
- اجري تقنين الاختبار على حوالي (1000) طفل و(400) يافع (مراحل البلوغ الأولي) اعتبروا من خلال الإجراءات الإحصائية عينة ممثلة للمجتمع الأمريكي.
- في النصف الثاني من العشرينات وأوائل الثلاثينات ظهرت الحاجة لإعادة بناء

هذا الاختبار من جديد، حيث وجدت الدراسات المستمرة لنتائج تطبيق واستخدام هذا الاختبار أن بعض اختباراتها الفرعية أصبحت تتمتع بصدق ضئيل، وان التعليمات لاستخدامه وتدريجه تعوزها الدقة والضبط، كما إنه بحاجة إلى صيغة بديلة مكافئة له والى مراجعة ادق للعينة التي اعتبرت ممثلة للمجتمع الأمريكي حينها.

- من كل ذلك كانت الحاجة إلى التنقيح الثاني الذي ظهر في عام 1937.

(ب) الطبعة الثانية 1937:

قام بها كل من لويس تيرمان ومود ميرل.

- قامت الطبعة الجديدة بناءاً على ما أسفرت عنه نتائج استخدام الاختبار في طبعته الأولى (كما ذكر في الفقرة (2) قبل قليل) إضافة إلى ما قدمته نتائج السنوات العشر المتتالية من البحث في مشروع التقنين الجديد لهذا الاختبار.
- قامت الطبعة الجديدة على ذات الخصائص المميزة لاختبار بينيه، متبعة ذات الافتراضات والطرق والمبادئ التي احتواه. لقد استخدمت المعايير العمرية للانجاز والافتراض القائل بأن الذكاء العام سمة تنمو مع العمر.
- إلا أن الطبعة الجديدة قامت بتقديم نماذج أكثر لذات أنماط الفعاليات العقلية مع احتفاظه بأساسية كونه مقياسا للقدرة العامة وليس لمجموعات القدرات الخاصة. كما أن الاختبار الجديد استمر بإعطائه وزنا اكبر للقدرة اللفظية كعنصر أساسي فيه مع استبعاده للعناصر التي لم تظهر ارتباطا جيدا أو أياً منه بالدرجة العامة التي يحددها الاختبار.
- قامت الطبعة الجديدة على أساس الصيغتين المتكافئتين للمقياس دعيت احدهما بصيغة (L) والأخرى بصيغة (M). واحتوت كل صيغة منه على (129) فقرة اختبارية تميزت بعدم التجانس الكبير في أنماطها المختلفة من ناحية وبعلاقتها

الوثيقة بالتقدير العام من ناحية بما ينسجم ويلائم المفهوم الأتي (وقته) المطروح حول القدرة العقلية العامة.

- اختيرت الفقرات الاختبارية (التي احتواها الاختبار في صيغته الجديدة بشقيه) من مئات الفقرات الاختبارية التي وضعت أساسا لبناء الاختبار منها. وقد كانت أسس اختيار الفقرات هو اختيارها إذا:
  - ارتبطت عاليا بالمعيار المقبول عن الذكاء.
  - ترتفع نسبة الإجابة عنها بتقدم العمر.
  - حصلت على النسبة المميزة بين من ينجح ويفشل فيها.
- اختيرت وفق المعايير اعلاه (209) فقرات اختيارية للصيغة (L) و(199) فقرة للصيغة (M) ثم طبقت هذه الفقرات على (3184) مفردة تجريبية. وبعد ذلك أجريت عليها عملية الغربلة الاحصائية من خلال عمليتين لتحقيق صدقها.

1- الزيادة أو التصاعد في الاستجابة الصحيحة لها من قبل الأفراد بزيادة أعمارهم.

2- ارتفاع الارتباط المتسلسل الثنائي Biserial Correlation لكل فقرة اختبارية مع الدرجة العامة التي يحددها الاختبار بكليته للمفردة التجريبية.

- أما ما يتعلق بالعينة فقد كانت عينة ممثلة خاصة بكل فئة عمرية عينة عشوائية طبقية)

- كانت الـ (3184) مفردة تجريبية من المواليد البيض الأصليين من سكان الولايات المتحدة الأمريكية.

- كانت هناك (100) مفردة تقريبا لكل فاصلة عمرية نصف سنوية مابين (1,5 إلى 5، 5) من العمر، و(200) مفردة لكل فئة عمرية سنوية مابين (6-14) من العمر و(100) مفردة لكل فئة عمرية سنوية مابين (15-18) من

العمر مع احتواء كل هذه الفئات العمرية بعيناتها على (50٪) من الذكور ومثلها من الإناث.

- عني بالعينة الجديدة التوزيع الجغرافي لها حيث احتوت على (17) جالية Communities في (11) ولاية متباعدة ضمت في كل منها مراكز المدن، أطرافها والقرويين مع الاحتفاظ بضبط العملية العشوائية للاختبار في كل مجتمع (أو جالية) من مجتمعات الدراسة.

كما كان متوسط المستوى الاقتصادي - الاجتماعي للعينة أكثر قليلا من المتوسط العام للمجتمع الأمريكي كما أعطته الإحصاءات المختلفة ذلك الوقت.

أجرى الثبات على الاختبار من خلال الارتباط ما بين النتائج التي قدمتها الصيغتان (L,M) بفرق أسبوع واحد بين الاستخدامين، وقد أعطت النتائج درجة عالية من الثبات وقد درست العلاقات التالية الخاصة بسبب التطبيقات الإكلينيكية لها:

- الثبات والعمر.
- الثبات ومستوى نسب الذكاء (I Q) للأفراد.

وقد اظهر تحليل ماكينمار Nc Nemar للبيانات المعيارية أن الاختبار أكثر ثباتا عند اليافعين منه عند الأطفال، وهو أكثر ثباتا عند اصحاب المستويات الواطئة من نسب الذكاء منه عند أصحاب المستويات العالية منها، فعلى سبيل المثال:

- مابين عمر(2,5 - 5,5): كان معامل الثبات 0,83 لفئة (140-149) من نسب الذكاء 0,91 يساوي (60-69) من نسب الذكاء.
- مابين عمر 6-13: كان معامل الثبت 0,91 = (140-149) من نسب الذكاء.
- مابين عمر 14-18: كان معامل الثبات 0,95 = (140-149) من نسب الذكاء 0,98 يساوي (60-69) من نسب الذكاء.

(ج) الطبعة الثالثة (1960):

- جمعت الصيغتان (L,M) التي احتوتهما الطبعة الثانية (1937) في صيغة واحدة سميت (L-M).

- كانت الغاية الرئيسية للتنقيح الجديد هي إعادة المعايرة (التقنين) للفقرات الاختبارية بعدما أوضحت التطبيقات الواسعة التي أجريت عليها في الخمسينات أن نسب المجيبين عن هذه الفقرات في الفئات العمرية ذات العلاقة قد تغيرت كثيرا عما كانت عليها عند استخدامها في الثلاثينات.

مثلا: في الاختبار الخاص للفئة العمرية (3) سنوات كان هناك (69٪) من الأطفال خلال الثلاثينات يستطيعون تمييز خمسة أشياء من (6) بصورة صحيحة في اختبار الرسوم الصغيرة التي تعرض عليهم للتعرف على (الحذاء، الساعة، التلفون، العلم، السكينة، الموقد) بينما لم يكن هناك غير (11٪) فقط من أطفال الخمسينات من يستطيع التعرف عليها بذات الطريقة.

- في الاختبار الخاص بالفئة العمرية (14) سنة كان معرفة الخطأ في صورة (الظل) أسهل عند شباب الخمسينات بنسبة (80٪) مما هي عند شباب الثلاثينات (63٪).

- يمكن أن تجمع أهم الملامح الرئيسية للطبعة الجديدة في كونها ذات صيغة واحدة وذلك منعا للازدواجية في الفقرات الاختبارية. إذ أن تحديد الفقرات في صيغة واحدة جعل من الممكن اختبار أكثر وأفضل الفقرات تمييزا في الصيغتين دونما تضحية بالإجراءات الأساسية لبناء الاختبار. كما أن الحاجة للصيغة البديلة أصبحت اقل بكثير في الخمسينات عما كانت عليه في الثلاثينات بعدما ظهرت اختبارات ذكاء أخرى خاصة بالأطفال الأمريكيين (اختبارات وكسلر مثلا) كما أن الدراسات أشارت إلى أن الثلاثينات والأربعينات شهدت استخدامها للصيغة الأولى (L) أكثر بكثير (بخمس أضعاف) من استخدام الصيغة الثانية (M) حيث كاد يكتفي الباحثون والأخصائيون بصيغة واحدة فقط.

- أما المواد ومصادر الفقرات الاختبارية الجديدة، فقد أجريت في الطبعة الجديدة على أساس الاستخدام التجريبي لها الذي امتد للسنوات (1950–1954) وإجراء المقارنة بين النتائج الجديدة أعطتها استخداماتها في أوائل الثلاثينات للمعايرة.
- كانت العينة الجديدة متألفة من (4498) مفردة تجريبية مابين 2,5–18 سنة من العمر. اختيرت عشوائيا من المجتمع العام دون تحديد لفئة عمرية معينة أو سحب عينة عشوائية طبقية للفئات العمرية كما جرى ذلك عند تنقيح الطبعة الثانية.
- يمكن أن يقال أن أهم التغيرات التي حدثت مابين طبعتي 1937 و 1960 يمكن تصنيفها بالنسبة إلى كل من المحتوى والتركيب.
- فمن حيث المحتوى اختيرت الفقرات ذات الكفاءة الأعلى من حيث معاملات الصعوبة والتمييز، حيث أبعدت الفقرات التي كانت تعطى ازدواجية في التطبيق. كما اختيرت الفقرات التي كانت أوضح وأيسر عند استخدامها وتدرجها.
- إن التغيرات في مستوى صعوبة اختبارات هذا المقياس قد حددت عن طريق المقارنة بنسب من تجاوزها (أي نجح فيها) في الثلاثينات متضمنة المجموعة المعيارية الاصلية. كما كان المحك لانتقاء العناصر يتم وفق معيارين:
  1- الزيادة في نسب الإجابات الصحيحة على الفقرة بتقدم العمر.
  2- الصدق المحدد عن طريق استخدام الارتباط المتسلسل الثنائي للفقرة الاختبارية بالدرجة العامة للاختبار.
- أما من حيث التركيب أو البنية الأساسية للاختبار فقد كانت التطبيقات الأخيرة للطبعة 1937 تشير إلى أن المتوسط بدأ يرتفع قليلا عن الـ (100) مما يستوجب أن تزداد صعوبة بعض الفقرات الاختبارية لتنخفض إلى الـ (100). كان هناك تصحيحان أساسيان:
  1- تعديل نسب الذكاء للتغاير النموذجي.
  2- تحديد جداول نسب الذكاء لتحتوي على الفئتين العمريتين (17، 18) لمسايرة

الافتراض الجديد القائل بأن النمو العقلي يستمر إلى ما بعد سن السادسة عشرة من العمر.

- ولقد أعطى التقنين الأخير للاختبار متوسطا مقداره (100) لكل فئة عمرية بانحراف معياري مقداره (16) درجة من درجات نسب الذكاء.
- كما أجريت بعض التعديلات (لأجل الوضوح والتيسير) في إجراءات استخدام الاختبار وفي طريقة تصحيح الإجابات.

وفيما يلي وصفا لكل مقياس من المقاييس الثلاث:

مقياس وكسلر لذكاء الكبار WAIS

يتألف مقياس وكسلر لذكاء الكبار من قسمين رئيسيين هما القسم اللفظي والقسم الأدائي ولا يتضمن المقياس اختبارات بديلة وتعطى الاختبارات بالترتيب التالي:

القسم اللفظي: ويتضمن من الاختبارات الفرعية الستة التالية:

- اختبار المعلومات information test
- اختبار الاستيعاب comprehension test
- اختبار الحساب arithmetic test
- اختبار المتشابهات similarities test
- اختبار إعادة الأرقام digit span test
- اختبار المفردات vocabulary test

القسم الأدائي: ويتضمن الاختبارات الفرعية الآتية:

- اختبار الترميز digit symbol، coding test
- اختبار تكميل الصور picture completion test
- اختبار تصميم المكعبات blook design test

- اختبار ترتيب الصور picture arrangement test
- اختبار تجميع الاشياء object assembly test

مقياس وكسلر لذكاء الأطفال WAIS

بني عام 1949 وتم مراجعته عام 1974 ويتألف مقياس وكسلر لذكاء الأطفال من 326 فقرة موزعة على الاختبارات الفرعية المكونة من القسم اللفظي والأدائي. حيث يتكون كل اختبار فرعي من عدد من الفقرات. ومن المقاييس الفرعية هي:

القسم اللفظي

- المعلومات العامة
- الاستيعاب
- الحساب
- المتضادات والمتشابهات
- المفردات
- إعادة الأرقام. (يبلغ مجموع فقرات القسم اللفظي 133)

القسم الأدائي

- تكميل الصور
- ترتيب الصور
- تصميم المكعبات
- تجميع الأشياء
- الترميز
- المتاهات. (يبلغ مجموع فقرات القسم الأدائي 193)

مقياس وكسلر لذكاء أطفال ما قبل المدرسة

Wechsler Preschool and Primary scale of intelligence

من تأليف Wechsler. D 1963 ويطبق على الفئة العمرية 4-6,6 سنة. يتكون المقياس من جزأين لفظي وأدائي ويتألف الجزء اللفظي من خمسة اختبارات فرعية أساسية وهي: المعلومات، المفردات، الحساب، المتشابهات، الاستيعاب. كما يتضمن اختبار احتياطياً وهو اختبار الجمل. أما الجزء إلا دائي فيتألف من خمسة اختبارات فرعية هي: بيت الحيوان، تكميل الصور، المتاهات، الأشكال الهندسية، تصميم المكعبات. يستخدم المقياس للأغراض التشخيصية الإكلينيكية وتقدير مستوى الأداء العقلي العام للأطفال ويطبق الاختبار فردياً ويستغرق تطبيقه ساعة واحدة.

قنن الاختبار على(1200) طفل من الجنسين وحسبت معاملات الثبات بالطريقة النصفية لكل اختبار فرعي وللجانب اللفظي والأدائي وتراوحت قيم المعاملات للدرجات اللفظية والأدائية ما بين (0,91-0,97) وللاختبارات الفرعية بين (0,62-0,91).

واستخرجت دلالات الصدق عن طريق صدق المحك بدراسة معاملات الارتباط بين اختبار وكسلر واختبارات ذكاء أخرى مثل اختبار ستانفورد بينيه وبلغ معامل الارتباط بينهما 0,91.

مقياس وكسلر لذكاء الأطفال (الصورة الاردنية)

أجرى القريوتي عام 1981 دراسة هدفت إلى تطوير صورة معربة ومعدلة من مقياس وكسلر لذكاء الأطفال WISC 1949 تتحقق فيها دلالات الصدق والثبات وفاعلية الفقرات. وقد أعدت الصورة الأولية للمقياس وتم تجريبها في ثلاث فئات عمريه وهي (5-7) (5-10) (5-13). طبقت الصورة الأولية من المقياس على 120 مفحوصاً بهدف التوصل إلى الصورة المعدلة. استخرجت دلالات الصدق من المقياس وذلك من خلال معاملات الترابط بين الأداء على المقياس في القسم اللفظي والأدائي.

كما توفرت دلالات الصدق من خلال معاملات الارتباط بين الأداء على المقياس والأداء على الاختبارات التحصيلية (اللغة العربية،العلوم،الرياضيات) في الفئات العمرية (5-10) (5-13). وتراوحت معاملات الارتباط ما بين 0,40 –0,61 كما تم استخراج دلالات الثبات باستخدام الطريقة النصفية وباستخدام طريقة الاعادة وتراوحت معاملات الثبات للقسم اللفظي ما بين 0,85-0,91 وللقسم الادائي 0,87 وللمقياس الكلي 0,93-0,94 أما معاملات الثبات المحسوبة بطريقة الاعادة وتراوحت معاملات الثبات للقسم اللفظي ما بين 0,85-0,91وللقسم الادائي0,87 وللمقياس الكلي 0,93-0,94 أما معاملات الثبات المحسوبة بطريقة الاعادة فتراوحت ما بين 0,78-0,80.

الصورة الأردنية من مقياس وكسلر لذكاء الأطفال مراجعة عام 1974

أجرى عليان والكيلاني عام 1988 دراسة هدفت إلى تطوير صورة أردنية معدلة من مقياس وكسلر لذكاء الأطفال –مراجعة عام 1974 تتوفر فيها دلالات صدق وثبات وفاعلية الفقرات في البيئة الأردنية. وعدلت الصورة الأولية للمقياس وجربت على 110 طفلاً وللفئات العمرية من (5-6) (5-16) وغطت ثلاث مناطق أردنية. وتوفرت للمقياس دلالات عن الصدق العاملي في صورته الأردنية كما توفرت دلالات عن الصدق التمييزي بين الفئات العمرية وذلك باتباع أسلوب تحليل التباين المتعدد للمتغيرات التابعة بين الفئات العمرية الخمسة. وحسب الثبات بالطريقة النصفية المصححة وبطريقة الإعادة فقد كانت معاملات الثبات للقسم اللفظي 0,92 والقسم إلا دائي 0,91 والمقياس الكلي 0,94

تألف مقياس وكسلر لذكاء الأطفال مراجعة 1974 من قسمين هما القسم اللفظي وعدد فقراته 132 موزعة على الاختبارات الفرعية:

- اختبار المعلومات ويتكون من 30 فقرة
- اختبار المتشابهات ويتكون من 17 فقرة

- اختبار الحساب ويتكون من 18 فقرة
- اختبار المفردات ويتكون من 32 فقرة
- اختبار الاستيعاب ويتكون من 17 فقرة
- اختبار إعادة الأرقام ويتكون من 14 فقرة

أما القسم إلا دائي من المقياس فيتضمن 200 فقرة موزعة على الاختبارات الفرعية التالية:

- اختبار تكميل الصور ويتكون من 26 فقرة
- اختبار ترتيب الصور ويتكون من 12 فقرة
- اختبار تصميم المكعبات ويتكون من 11 فقرة
- اختبار تجميع الأشياء ويتكون من 4 فقرات
- اختبار الترميز ويتكون من 138 فقرة
- اختبار المتاهات ويتكون من 9 فقرات

مقياس فروستنج للإدراك البصري

بني مقياس فروستنج للإدراك البصري المعروف باسم frosting developmental test of visual perception 1961 وتم مراجعته عام 1966. يهدف المقياس إلى قياس وتشخيص مظاهر الإدراك البصري في الفئات العمرية من (3–8) سنوات. ويعتبر هذا المقياس من المقاييس الفردية-الجمعية المقننة المشهورة في مجال الإدراك البصري وخاصة للأطفال ذوي الإعاقة البصرية الجزئية وذوي صعوبات التعلم.

يتألف المقياس من (57) فقرة موزعة على الاختبارات الفرعية الآتية:

1- مقياس التآزر البصري الحركي ويتألف من 16 فقرة.

2- مقياس التمييز بين الشكل والأرضية ويتألف من 8 فقرات.

3- مقياس ثبات الأشكال ويتألف من 17 فقرة.

اختبار كاليفورنيا للنضج العقلي

California short-from test of mental maturity

بني الاختبار من قبل Sullivan E. T.، clarte w. w. & tiges E. W ويطبق على الفئات العمرية من(10) سنوات وحتى سن الرشد. صنف الاختبار القدرات العقلية في أربعة مجالات وهي: العلاقات المكانية، المحاكمة المنطقية، المحاكمة العددية، المفاهيم اللفظية.

يزودنا الاختبار بمعلومات عن طبيعة القدرات العقلية وبنائها لدى الأفراد. ويتكون الاختبار من سبعة اختبارات فرعية يمكن بناءاً عليها أن نجد نسبة التطور العقلي التي تقيس المجالات الأربعة السابقة ويضم مجالات العلاقات المكانية اختباراً في التعرف على اليد اليمنى واليد اليسرى واختباراً للتعامل مع الاتجاهات. أما مجال المحاكمات المنطقية فيضم المتشابهات والاستنتاج في حين يضم مجال المحاكمة العددية السلاسل الرقمية والمسائل العددية ويضم مجال المفاهيم اللفظية اختبار هذه المفاهيم.

يستخدم الاختبار كأداة للتعرف على مستوى التطور العقلي لدى الأفراد ويوفر المقياس معلومات أساسية يمكن أن تستخدم أساساً لتفسير مستوى أداء الفرد الحالي والاستعداد المستقبلي لديه لممارسة بعض النشاطات الأساسية المحددة. ويوفر الاختبار معلومات مهمة بشكل خاص للمعلمين والمرشدين وعلماء النفس وأرباب العمل وتستخدم أيضاً لأغراض التوجيه المهني وتساعد على تكييف موقف التعليم القدرة لدى الفرد إلى أقصى درجة ممكنة. يطبق الاختبار فردياً وجمعياً ويستغرق وقت التطبيق (52) دقيقة.

قنن الاختبار على عينة مكونة من(2500) حالة اختيرت بشكل عشوائي طبقي يغطي المناطق الجغرافية في الولايات المتحدة الأمريكية للمستويات التعليمية المختلفة،وجرى التقنين على عينة من الأسوياء وغير الأسوياء من طلاب المدارس.

حسب الثبات عن طريق الاتساق الداخلي بتطبيق معادلة كورد ريتشاردسون حيث بلغت للجزء اللفظي 85-0 وللجزء غير اللفظي 0,96 وللاختبار ككل 0,88 كما حسبت بطريقة التجزئة النصفية فكان معامل الارتباط الكلي 94، 0 وللجزء غير اللفظي 0,87 والجزء اللفظي 0,94

واستخرج الصدق بالتحليل العاملي فوجد أن الاختبار يقيس خمسة عوامل وعن طريق المحك فتراوحت معاملات الارتباط بين هذا الاختبار واختبار وكسلر بين 0,65-0,93

اختبار أوتس-لينون للقدرة المدرسية Lennon school ability test Otis

بني من قبل Arthurs. Otis & Roger T. lennon عام 1979 وتم مراجعته عام 1982 وللمراحل الدراسية من الأول إلى المرحلة الثانية عشر.

ويستخدم لتزويد المعلمين بالمعلومات عن قدرات الطلبة ويساعدهم في وضع الخطط والبرامج التربوية الملائمة ويتكون من المستويات الآتية:

1- المرحلة الابتدائية الأولى وللمرحلة الأولى ويستغرق وقت تطبيقه (80) دقيقة
2- المرحلة الابتدائية الثانية وللمراحل الثانية والثالثة ويستغرق وقت تطبيقه (80) دقيقة.
3- المرحلة المتوسطة وللمراحل الرابعة والخامسة ويستغرق الوقت (45) دقيقة
4- المرحلة الثانوية وللمراحل السادسة والثامنة ويستغرق الوقت (45) دقيقة..
5- المرحلة المتقدمة وللمراحل التاسعة والثانية عشر ويستغرق الوقت (40) دقيقة.

يقيس الاختبار ثلاث قدرات أساسية وهي: القدرة اللفظية والقدرة العددية والقدرة على الاستدلال المجرد وتقاس هذه القدرات باستخدام أنواع متعددة من الفقرات في كل مستوى من المستويات ومن الأمثلة على أنواع القدرات المتضمنة هي:

- المعلومات العامة
- المتشابهات والمتضادات
- المترادفات
- الجمل المغلوطة
- السلاسل العددية
- الاستدلال الحسابي

ويستخدم الاختبار للكشف عن القدرات العقلية لدى الأفراد ويمكن استخدامه لغايات التشخيص والعلاج والإرشاد ولغايات تربوية أهمها التنبؤ بالنجاح المدرسي والجوانب المهنية التي تحتاج للقدرة على معالجة المفاهيم الرمزية.

وتم تحضير الفقرات من خلال التعريف للقدرة العقلية أولاً ثم وضعت الفقرات التي تمثل أبعاد القدرة العقلية للصفوف المختلفة باختلاف العمر. واتبع ذلك بعملية تحليل الفقرة واستبعدت الفقرات التي لا تتمتع بدرجة تمييز وصعوبة مناسبة ثم وضع نموذج مكافئ لكل مستوى حسب الصف بعد ذلك فحص مستوى القدرة على قراءة هذه الفقرات للمفحوصين من أعمار مختلفة وبين التحليل أن الاختبار لا يتطلب قدرات قرائية عالية. ويمكن تطبيق الاختبار بصورة جمعية. قنن الاختبار على عينة متكونة من(1229) مفحوص يمثلون مستويات دراسية واقتصادية واجتماعية مختلفة واستخرجت الرتب المئينية والمكافئات العقلية لكل عمر عقلي.

استخرج معامل الثبات عن طريق الصور المكافئة للاختبار وتراوحت قيمها ما بين (83،-0,95) للأعمار المختلفة وبالطريقة النصفية فتراوحت ما بين(0,88-0,95)أما طريقة إعادة الاختبار فتراوحت ما بين (0,80 –0,94).

واستخرج معامل الصدق عن طريق صدق المحتوى والتحليل العاملي للاختبارات الفرعية وعن طريق صدق المحك مع اختبارات أخرى مثل اختبار أوتس للمهارات

الأساسية وتراوحت قيم معاملات الارتباط ما بين (0,85-0,71) للاختبارات الفرعية في أعمار مختلفة وفحصت معاملات الارتباط بين درجة الاختبار ودرجات الاختبارات التحصيلية فتراوحت قيمها بين (0,78-0,54) وعن طريق معاملات الارتباط بين درجات الاختبار واختبار بينية للذكاء وتراوحت (0,83-0,42) وحسبت المعايير المئينية للاختبار.

مقياس الينوي للقدرات السيكولغوية

بني مقياس الينوي للقدرات السيكولغوية من قبل kirk,mccarthy & kirk 1961 وتم مراجعته عام 1968 ويهدف المقياس إلى قياس وتشخيص مظاهر الاستقبال والتعبير اللغوي وخاصة لدى الأطفال ذوي صعوبات التعلم ويعتبر المقياس من المقاييس الفردية المقننة والمشهورة في مجال صعوبات التعلم.

يتألف المقياس من 12 اختبار فرعياً تغطي طرق الاتصال اللغوية ومستوياتها والعمليات النفسية والعقلية التي تتضمنها تلك الطرق والاختبارات الفرعية هي:

1- اختبار الاستقبال السمعي Auditory Reception subtest ويقيس قدرة الطفل على الاستقبال السمعي والإجابة بكلمة نعم أو لا.

2- اختبار الاستقبال البصري Visual Reception subtest ويقيس الاختبار قدرة الطفل على مطابقة صور مفهوم ما مع صورة أخرى ذات علاقة.

3- اختبار الترابط السمعي Auditory Association subtest ويقيس الاختبار قدرة الطفل على إكمال جمل متجانسة في تركيبها اللغوي.

4- اختبار الترابط البصري Visual Association subtest ويقيس الاختبار قدرة الطفل على الربط بين المثيرات البصرية المتجانسة أو ذات العلاقة.

5- اختبار التعبير اللفظي Verbal Expression subtest ويقيس الاختبار قدرة الطفل على التعبير اللفظي عن الأشياء التي يطلب منع تفسيرها.

6- اختبار التعبير العملي Manual Expression subtest ويقيس قدرة الطفل على التعبير عملياً أو يدوياً عما يمكن أداؤه بأشياء معينة.

7- اختبار الإكمال القواعدي Grammatic Closure subtest ويقيس الاختبار قدرة الطفل على إكمال جمل ذات قواعد لغوية مترابطة.

8- اختبار الإكمال البصري Visual Closure subtest ويقيس الاختبار قدرة الطفل على إدراك وتمييز موضوعات ناقصة حيث تعرض على المفحوص لوحة تتضمن عددا من الموضوعات الناقصة ويطلب منه تمييزها.

9- اختبار التذكر السمعي Auditory Memory subtest ويقيس الاختبار قدرة الطفل على تذكر سلاسل من الأرقام تصل في أقصى مدى لها إلى 8 أرقام حيث تطرح على المفحوص بمعدل رقمين في كل ثانية.

10- اختبار التذكر البصريVisual Memory subtest ويقيس الاختبار قدرة الطفل على تذكر أشكال لا معنى لها بطريقة متسلسلة حيث يعرض على المفحوص كل شكل من تلك الأشكال لمدة خمس ثواني ويصل أقصى مدى لتلك الأشكال إلى ثمانية أشكال.

11- اختبار الإكمال السمعي Auditory Closure subtest وهو اختبار احتياطي ويقيس قدرة الطفل على إكمال مفردات ناقصة متدرجة في مستوى صعوبتها.

12- اختبار التركيب الصوتي Sound Blending subtest وهو اختبار احتياطي ويقيس قدرة الطفل على تركيب الأدوات معاً حيث يطلب من المفحوص أن يركب الأصوات التي يسمعها بفاصل زمني قدره نصف ثانية بين كل صوت وآخر حيث يبدأ الفاحص بعرض أصوات الكلمات ذات معنى ثم ينتهي الفاحص بعرض أصوات لكلمات لا معنى لها.

يصلح الاختبار للأطفال في الفئة العمرية من 2-10 سنوات ويستغرق تطبيقه حوالي ساعة ونصف. واستخرجت دلالات الصدق للمقياس في صدق البناء

والمفهوم والصدق التلازمي كما استخرج من المقياس العوامل التي تفسر الأداء على المقياس باستخدام أسلوب التحليل العاملي.

واستخرج الصدق التلازمي للمقياس من خلال الترابط بين المقياس الحالي ومقاييس أخرى كالاختبار التحصيلي الواسع المدى واختبار كاليفورنيا للتحصيل ومقياس ستانفورد بينية للذكاء ومقياس وكسلر وتراوحت معاملات الارتباط ما بين 0,49-0,51.

كما ظهرت دراسات تؤكد الصدق التنبؤي من خلال الترابط ما بين المقياس الحالي ومقياس كاليفورنيا للتحصيل فيما بعد وبلغ 0,40 كما توفرت دلالات عن ثبات المقياس إذ بلغ معامل الثبات 0,45 بطريقة الاتساق الداخلي لاختبار الإكمال السمعي و0,96 لاختبار الاستقبال السمعي. وقنن المقياس على عينة مؤلفة من 962 طفلا في بعض من الولايات المتحدة الأمريكية.

اختبارات الاستعدادات الفارقة The Differential Aptitude DAT

لقد أعدها Bennett بنييت وseashore شيور wesman عام 1947، وقد عدلت عام 1963 وأعيد تقننيها وظهرت الطبعة الرابعة لها عام 1966 وتصلح بصورة خاصة للمستويات العمرية من (14-18) سنة وتستخدم في أغراض التوجيه التربوي والمهني وتتألف البطارية من (8) اختبارات وتقاس القدرات المتضمنة في هذه البطارية قياسا مستقلاً وهي الاستدلال اللفظي والقدرة العددية والتفكير المجرد والعلاقات المكانية والاستدلال الميكانيكي والسرعة والدقة الكتابية والاستخدام اللغوي بقسميه: القدرة على تهجي الكلمات والقدرة على استخدام القواعد اللغوي.

وللبطارية صورتان متكافئتان L, M وقد طبع كل اختبار في كتيب منفصل تيسراً للاستفادة العلمية. وقد تم تقنين البطارية على (47) ألف تلميذ يمثلون ستة وعشرون ولاية أمريكية، وتم حساب المعايير للإناث وأخرى للذكور في كل سنة

دراسية، وكانت الدرجات الخام تحول إلى درجات مئينية، كما يمكن قراءة الدرجات المعيارية مباشرة من لوحة البروفيل المئيني.

اختبارات فلانجان لتصنيف الاستعدادات F A C T

ظهرت هذه البطارية نتيجة لبحوث(Flanagan)حول أعداد اختبارات تصنيف القوات الجوية بالجيش الأمريكي أثناء الحرب العالمية الثانية وتهتم بمجال التوجيه المهني وقد اعتمدت على منهج تحليل العملJJob analysis لكثير من المهن. وتتألف البطارية في صورتها الأولية من (21) اختبار.

وقد عدلت البطارية فأصبحت تشمل (19) اختباراً وتقسم الاستعداد إلى (32) وظيفة مختلفة. والاختبارات هي:

- الفحص: القدرة على كشف الأخطاء
- الترميز: السرعة والدقة في ترميز المعلومات المكتبية الطباعية
- الحفظ: تذكر تعلم الترميز في اختبارات الترميز
- الدقة: السرعة والدقة في عمل حركات دائرية صغيرة للإصبع
- التجميع: القدرة على تكوين صور ذهنية لشيء معين حين توضع مجموعة من الأجزاء معا "
- المقاييس: السرعة والدقة في قراءة المقاييس والرموز البيانية
- التآزر: القدرة على تناسق حركات اليد والساعد
- الحكم والفهم: القدرة على الفهم والقراءة والتعليل واستخدام الحكم الجيد
- الحساب: العمليات الأساسية الأربع
- الأنماط: القدرة على إنشاء نسخ المخططات البسيطة
- المكونات: القدرة على تحديد أجزاء المكونات المهمة في الرسوم أو الرسوم التخطيطية

- الجداول: قراءة نمطين من الجداول، الأول استخدام الأرقام والآخر استخدام الكلمات
- الميكانيكية: فهم الأعمال الميكانيكية ومبادئها
- التعبير: اتصال الأفكار في التحدث والكتابة
- الحذق: القدرة على الابتكار والإبداع والاختراع في إعداد إجراءات أو الآلات أو عروض ذكية اختبار استدلالي يعتمد على مشكلة محددة
- اليقظة: القدرة على تقدير الموقف وملاحظة ما ينشأ فيه من أخطار أو ما يحتاجه من وسائل وخطط.
- الاستدلال: القدرة على حل المشكلات والتعبير عن الحل بصورة رياضية
- المفردات: القدرة على فهم معاني الكلمات
- التخطيط: الكفاءة في تنظيم البيانات بإعادة ترتيب خطوات معينة لحل المشكلة

اختبارات ايوا للنمو التعليمي Iwa tests of Education Development

وضعت هذه الاختبارات لقياس المهارات الأساسية لتلاميذ الصفوف ابتداءً من (السنة الثالثة للمرحلة الابتدائية وحتى السنة الثانية للمرحلة المتوسطة) وتتضمن (9) اختبارات هي: فهم المفاهيم الاجتماعية الأساسية، وصحة الكتابة، والمعلومات العامة في العلوم الطبيعية والقدرة على التفكير الكيفي والقدرة على تغيير البيانات المقروءة في العلوم الاجتماعية، وفي العلوم الطبيعية، والقدرة على تفسير المواد الأدبية والمفردات العامة واستخدام مصادر المعلومات. وقد حدث تطوير لهذه الاختبارات كي تلائم أغراض الاختبار والتوجيه في سلاح الطيران الأمريكي ولذلك أصبح اسمه اختبار النمو العقلي العام لمعهد سلاح الطيران الأمريكي.

اختبارات المتروبوليتان التحصيلية: (M. A. T)

وهي من أشيع بطاريات التحصيل استخدامها في مستوى المتوسطة وان كانت

تستخدم في المدرسة الابتدائية ابتداءً من الصف الأول وحتى نهاية المرحلة المتوسطة. وتتكون من (4) اختبارات وهي: صورة الكلمة، التعرف على الكلمة، معنى الكلمة والأرقام. كما ظهرت اختبارات المتروبوليتان للتهيؤ في القراءة التي تتضمن (6) اختبارات فرعية هي معنى الكلمة، الجمل، المعلومات، المزاوجة، الأعداد، النسخ.

اختبارات كيتس للتهيؤ في القراءة: Gates Reading Readiness tests:

وتتكون من (5) اختبارات فرعية لقياس التهيؤ لتعلم القراءة والتنبؤ بسرعة نمو القدرات القرائية ولتشخيص مكانة التلميذ وبيان حاجته في كل قدرة هامة يحتاجها عند تعلم القراءة والاختبارات الخمسة هي: مزاوجة الكلمات، المزاوجة بين الكلمة وبطاقات تعليم تناغم الأصوات Rhyiming وقراءة الحروف، والأرقام.

اختبارات النمو العام للأطفال

ثمة عدد كبير من الاختبارات تقيس النمو العام للطفل منذ الولادة وحتى سن المدرسة وتركز على النمو الحركي واللغوي والمعرفي والانفعالي ومن أهمها:

1- مقاييس جيزل للنمو والذي يضع معايير عامة تتبع نمو سلوك الطفل من الأسبوع الأول وحتى(5) سنوات إذ تحسب التطورات في الأسابيع خلال السنة الأولى أو في الأشهر فيما بعد وحتى (60)شهراً وتعطي متوسطاً عاما للنمو الحركي واللغوي والتكيف العام بما في ذلك نموه الاجتماعي وقد وضع جيزل هذه الاختبارات عام 1919 وأجرى عليها دراسات وتعديلات قادت إلى تحسين وتطوير هذه المقاييس.

2- اختبارات بيلي وميرى والتي وضعت عام 1931 لقياس النمو الحركي واللغوي والعقلي للأطفال من شهر وحتى ثلاثة شهور.

3- اختبارات بياجيه والتي بدأ بها منذ الثلاثينيات، وهو أحد أهم علماء النفس

السويسريين والذي درس النمو المعرفي العقلي بمنهجية خاصة والذي افترض بأن النمو العقلي للأفراد جميعاً يمر بأربع مراحل أساسية اثنتان منها في مرحلة الطفولة ما قبل السن المدرسية، وان الاختلاف في ذكاء الأفراد يرجع إلى قدرة البيئة على مساعدة الفرد في الانتقال من مرحلة إلى أخرى وتتلخص هذه المراحل إلى:

أ- الإدراك الحسي-الحركي من الولادة وحتى السنة الثانية.

ب- التفكير الحدسي من 3-6 سنوات.

ج- التفكير المحسوس من 7-11.

د- التفكير الرمزي المجرد من 12 فما فوق.

وثمة تقسيمات داخلية لكل مرحلة من هذه المراحل وقد صممت لها اختبارات خاصة طبق البعض منها في العراق وفي اغلب بلدان العالم.

الاختبارات المتحررة ثقافيا

وهي اختبارات تقيس الذكاء العام بمعزل عن القدرة اللفظية أي أنها اختبارات غير لفظية تعتمد على الرسوم والأشكال التي لا ترتبط بأي ثقافة طبقية.

ومن هذه الاختبارات ما يأتي:

1- **اختبار رسم الرجل لكود أنف وهاريس** والذي مضى على فكرته الأساسية ما يقارب المئة عام وجرى عليه تطوير بشكله الاخير،حيث يطلب الشخص أن يرسم رجلاً والى البنت أن ترسم امرأة والافتراض النظري يعتمد على علاقة رسم الشخص بتفاصيله الكاملة مع درجة ذكاء الفرد، بغض النظر عن معرفة الشخص ومهارته في الرسم وقد جرى أكثر من محاولة لتطبيق هذا الاختبار في العراق، كما جرب في عدد من الدول فضلا عن شهرته في اوربا وأمريكا ليس لقياس الذكاء فحسب بل لمعرفة اضطرابات الشخصية.

2- **اختبار المصفوفات المتتابعة لرافن** والذي يتكون من ثلاثة مجاميع من الرسوم

تحتوي على 12 سؤالا على شكل مصفوفة لرسوم وأشكال ناقصة يطلب تكملتها من بدائل مصورة أسفل كل سؤال.

في ضوء الإجابات تحدد درجة الذكاء وقد جرب في العراق وعدد من الدول العربية الأخرى ويستعمل في تصنيف الأطفال والكشف عن المتخلفين عقلياً.

3- **اختبار الدومينو** وهو اختبار للذكاء العام يتكون من 36 فقرة على شكل رسوم لقطع الدومينو يطلب فيها معرفة التسلسل المنطقي المكمل للمصفوفة المصورة ويعطي درجة ذكاء عام وقد استعمل وما زال يستعمل في عدد من البلدان الأوربية منها الجيش الفرنسي، كما جرب في بعض الدول الأفريقية والشرق الأوسط.

الاختبارات المتخصصة

وتشمل هذه الاختبارات عدد كبير من الاختبارات تقيس قدرات خاصة أو مركبة مثل الإبداع والابتكار والقدرة الموسيقية أو التذوق الفني ومن بين أشهر هذه الاختبارات اختبار تورانس للقدرات الإبداعية واختبار جيلفورد وغيرها ونستعرض هنا البعض من هذه الاختبارات.

نشرت هذه الاختبارات عام 1966 وهي تتألف من مجموعتين من الاختبارات الفرعية تقيس عناصر التفكير الابتكاري والتي تحدد وفق إطارها النظري بأربعة وهي:

1- الطلاقة وهي عدد الأفكار التي يقدمها الشخص لإضافة أو استعمال أو إكمال صورة أو كلمة في مدة زمنية محددة.

2- المرونة وهي عدد الأصناف لتلك الإجابات

3- الأصالة وهي أعلى الدرجات لأفضل الإجابات

4- التطوير وهو مقدار التفاصيل المستخدمة في الإجابات.

وقد جرب هذا الاختبار بعد تعديلات خاصة في مصر والعراق وعددا من الدول العربية

اختبارات القدرة الموسيقية

وهي تسجيلات موسيقية تقيس مجالات أساسية في الموسيقى هي:

1- تمييز النغمات من حيث حدة الذبذبات الصوتية
2- تمييز شدة الصوت من حيث الارتفاع والانخفاض
3- تمييز الانسجام بين نغمتين مختلطتين
4- تمييز المسافات الزمنية بين النغمات
5- التوقيت أو النغم المنظم على الوحدة
6- تذكر النغمات المتشابهة

يستمع الشخص في هذه التسجيلات إلى أزواج من النغمات وعليه أن يؤشر في كل زوج من النغمات أيهما أكثر حدة أو شدة أو انسجاماً وهكذا وتعتبر الدرجة العالية هي قياس للقدرة الموسيقية.

الفصل العاشر

مقاييس الاتجاهات

يحتوي الفصل على:

- مقاييس الاتجاه
- تعريف الاتجاه
- أنواع الاتجاهات
- خصائص الاتجاه
- وظائف الاتجاه
- الفرق بين الاتجاه والعاطفة، الرأي العام، التعصب، القيم
- تكوين الاتجاهات
- المجالات والأغراض التي تستخدم فيها مقاييس الاتجاهات
- قياس الاتجاهات
- صدق مقاييس الاتجاهات
- طرق قياس الاتجاهات
- تحليل المحتوى
- نماذج من مقاييس الاتجاهات والقيم

مقاييس الاتجاهات

يعد المفكر الانكليزي هربرت سبنسر من أوائل علماء النفس الذين استخدموا اصطلاح الاتجاهات في الوصول إلى الأحكام الصحيحة في المسائل المثيرة للجدل. أما المفكر الأمريكي "جور دون البورت" فأعتبر مفهوم الاتجاهات من ابرز المفاهيم وأكثرها إلزاما في علم النفس الاجتماعي الأمريكي المعاصر فليس هناك اصطلاح واحد يفوقه في عدد مرات الظهور في الدراسات التجريبية. ويرجع البورت سبب شيوع هذا الاصطلاح إلى العوامل التالية.

1- إن هذا الاصطلاح لا ينتمي إلى أي من المدارس السيكولوجية التي كان يسود بينها النزاع. وهي مدرسة الغرائز السلوكية ومدرسة الكشتالت. وعليه فمن الطبيعي أن يتلقفه غالبية علماء النفس الذين كانوا يقفون خارج هذه المدارس.

2- إن هذا الاصطلاح يساعد المتبني له أن يتهرب من مواجهة مشكلة البيئة والوراثة التي كان الجدل حولها محتدما طوال العقدين الثالث والرابع من هذا القرن.

3- إن لهذا الاصطلاح قدر من المرونة يسمح باستخدامه في نطاق الفرد وعلى نطاق الجماعة. وقد استخدم فعلا في كل من هاتين الوجهتين مما جعله نقطة التقاء بين علماء النفس وعلماء الاجتماع وأتاحت بينهم الفرص للمناقشة والتعاون في البحث.

4- الرغبة الملحة لدى علماء النفس بوجه عام وخاصة في أمريكا في أن يتمكنوا من استخدام المقاييس في دراستهم فالقياس في أذهان الكثيرين هو الذي جعل البحث جديرا بأن يسمى بحثا علميا.

وفي هذا السياق ستتعرض بالدراسة والتحليل هذا الموضوع الذي اعتبره علماء الدراسات السلوكية والنفسية إنه من أهم مواضيع علم النفس الاجتماعي، بل لقد ذهب البعض إلى اعتباره إنه هو الميدان الوحيد لذلك العلم، ويستند

أصحاب هذه الآراء إلى أن جميع الظواهر النفسية الاجتماعية، بسيطة كانت أم مركبة، خاصة أو عامة، تخضع في أساسها لمحددات السلوك الإنساني الذي يواجهه ويسيطر عليه تركيب خاص يسمى الاتجاه النفسي Psychology attitude بالإضافة إلى أن القيم والاهتمامات تؤثر بشكل واضح وفعال على هذا السلوك.

تعريف الاتجاه

ظهرت العديد من التعاريف حول الاتجاهات إلا أن التعريف الذي كان من أكثر التعاريف انتشارا هو تعريف البورت 1935 فيعرف الاتجاه بأنه "حالة من الاستعداد أو التأهب العصبي والنفسي، تنتظم من خلال خبرة الشخص وتكون ذات تأثير توجيهي أو دينا مي على استجابة الفرد لجميع الموضوعات والمواقف التي تستثير هذه الاستجابة. ويفسر هذا التعريف على أن الاتجاه هو الحالة الوجدانية القائمة وراء رأى الشخص أو اعتقاده فيما يتعلق بموضوع معين، من حيث رفضه لهذا الموضوع أو قبوله ودرجة هذا الرفض أو القبول. ويمكن القول بأن كل شخص يحمل نوعين من الاتجاهات:

1- اتجاهات خاصة أو شخصية وهذه هي مجموعة اتجاهاته نحو أحداث حياته الخاصة وظروفها من حيث هي خاصة به.

2- اتجاهات عامة أو اجتماعية وهذه هي مجموعة اتجاهاته نحو الأحداث والموضوعات العامة في الحياة الاجتماعية.

أما ثيرستون Thurstone فيعرفه بأنه درجة الميل الإيجابي أو السلبي المرتبط ببعض الموضوعات السيكولوجية.

ويقصد ثيرستون بالموضوعات السيكولوجية أي رمز أو قضية أو شخص أو مؤسسة أو مثال أو فكرة وغير ذلك مما يختلف حوله الناس. فالاتجاه لا يكون إزاء الحقائق الثابتة المقررة وإنما هو دائما تجاه الموضوعات التي يمكن أن تكون موضوعات جدلية.

أما "نيوكمب" فيؤكد عنصر الدافع في مفهوم الاتجاه ويرى أن الاتجاه حالة من الاستعداد تثير الدافع ومن ثم فإن اتجاه الفرد نحو شيء ما يصبح عبارة عن استعداد للعمل والإدراك والتفكير والشعور، أي الاستعداد للاستجابة أي كان نوعها، ولكن الاتجاه ليس هو السلوك ذاته أو الاستجابة ذاتها لكنه الدافع الذي يمكن وراء السلوك.

أما انستازي فيعرف الاتجاه بأنه النزعة أو الميل للاستجابة المؤيدة أو غير المؤيدة نحو مجموعة معينة من المثيرات كالجماعات العنصرية أو التقاليد أو العادات أو المؤسسات. ومن هنا يتضح من التعريف بأن الاتجاهات لا يمكن ملاحظتها مباشرة ولكن نستدل عليها من خلال السلوك الصريح، اللفظي، وغير اللفظي.

وفي أكثر المصطلحات الموضوعية نقول بأن الاتجاه هو مفهوم ثابت نسبيا للاستجابة نحو الأشياء والأصناف والموضوعات التي تستثير هذه الاستجابة ومن خلال التطبيق العملي نلاحظ أن مصطلح الاتجاه غالبا ما يرتبط بالمثيرات الاجتماعية وفي ضوء ما تقدم يمكن أن نحدد تعريفا للاتجاه بما يلي:

"استعداد نفسي أو حالة عقلية - ثابتة نسبيا - يتكون عند الفرد نتيجة العوامل المختلفة المؤثرة في خبراته ويكون ذا تأثير توجيهي على استجابة الفرد نحو أفكار معينة أو حوادث أو مواضيع أو أشخاص أو أشياء ويرتبط هذا الموضوع عادة بشعور داخلي لدى الفرد ويتم قياسه إجرائيا من خلال الدرجة التي يحصل عليها الفرد، نتيجة استجابته على عبارات المقياس الذي اعد لهذا الغرض"

وواضح من التعاريف السابقة أن الاتجاهات تتكون من خلال الخبرات التي يمر بها الفرد فليس الاتجاه إذن سمة فطرية أو موروثة وإنما هو مكتسب بالخبرة والتقليد والمحاكاة.

ولعل من ابرز خصائص الاتجاه ما يلي:

- إنه حالة عقلية وعصبية ثابتة ثبوتا نسبيا أو دائمة دواما نسبيا، فالفرد الذي

يحمل اتجاها ايجابيا نحو القيم الدينية لا يتغير سلوكه من موقف لآخر ولكن هذا بطبيعة الحال لا يمنع من كون الاتجاه ديناميكي متغير في طبيعته ولكن هذا التغير يحدث على المدى الطويل، فالاتجاه يتغير نتيجة لما يقع على الفرد من مؤثرات مختلفة نتيجة لتفاعله مع البيئة المادية والاجتماعية والثقافية التي يعيش فيها.

- أنها مكتسبة ومتعلمة وليست موروثة.
- أنها ترتبط بمثيرات ومواقف اجتماعية ويشترك عدد من الأفراد أو الجماعات فيها.
- أنها تتعدد وتختلف حسب المثيرات التي ترتبط بها.
- أنها توضح وجود علاقته بين الفرد وموضوع الاتجاه.
- تتفاوت الاتجاهات في وضوحها وجلائها لدى الأفراد فمنها ما هو واضح صريح ومنها ما هو غامض مستتر.
- أن الاتجاه النفسي يقع دائما بين طرفين متقابلين احدهما موجب والآخر سالب هما التأييد المطلق والمعارضة المطلقة.
- من الاتجاهات ما يبقى قوياً على مر الزمن على الرغم مما يقابل الفرد من شواهد تدعو إلى التخلي عنه وإسقاطه. مثل هذه الاتجاهات، اتجاهات قوية، بخلاف الاتجاهات الضعيفة التي تتغير وتتحول تحت وطأة العناء والشدائد.

أنواع الاتجاهات

تصنف الاتجاهات النفسية إلى الأنواع التالية:

1- الاتجاه القوي: يبدو الاتجاه القوي في موقف الفرد من هدف الاتجاه موقفا حادا لا رفق فيه ولا هوادة فالذي يرى المنكر فيضعف ويثور ويحاول تحطيمه انما يفعل ذلك لان اتجاها قويا حادا يسيطر على نفسه.

2- الاتجاه الضعيف: يتمثل في وقوف الفرد موقفا ضعيفا رخوا خانعاً مستسلما فهو يفعل ذلك لأنه لا يشعر بشدة الاتجاه كما يشعر بها الفرد في الاتجاه القوي.

3- الاتجاه الموجب: هو الاتجاه الذي ينحو بالفرد نحو شيء ما (أي ايجابي)

4- الاتجاه السالب: هو الاتجاه الذي يجنح بالفرد بعيدا عن شيء آخر (أي سلبي)

5- الاتجاه العلني: هو الاتجاه الذي لا يجد الفرد حرجا في إظهاره والتحدث عنه أمام الآخرين.

6- الاتجاه السري: هو الاتجاه الذي يحاول الفرد إخفائه عن الآخرين ويحتفظ به في قرارة نفسه بل ينكره أحيانا حين يسأل عنه.

7- الاتجاه الجماعي: هو الاتجاه المشترك بين عدد كبير من الناس فإعجاب الناس بالأبطال اتجاه جماعي.

8- الاتجاه الفردي: هو الاتجاه الذي يميز فرداً عن آخر فإعجاب الإنسان بصديق له اتجاه فردي.

9- الاتجاه العام: هو الاتجاه الذي ينصب على الكليات وقد دلت الأبحاث التجريبية على وجود الاتجاهات العامة فأثبتت أن الاتجاهات السياسية تتسم بالصفة العمومية ويلاحظ أن الاتجاه العام هو أكثر شيوعا واستقرارا من الاتجاه النوعي.

10- الاتجاه النوعي: هو الاتجاه الذي ينصب على النواحي الذاتية وتسلك الاتجاهات النوعية مسلكا يخضع في جوهره لإطار الاتجاهات العامة وبذلك تعتمد الاتجاهات النوعية على العامة وتشتق دوافعها منها.

وظائف الاتجاهات

1- إن للاتجاهات الدور الرئيس في تنظيم العمليات الدافعية والانفعالية والإدراكية والمعرفية حول بعض النواحي الموجودة في المجال الذي يعيش فيه الفرد، حيث

ييسر للفرد القدرة على السلوك واتخاذ القرارات النفسية في شيء من الاتساق والتوحيد دون تردد أو تفكير في كل موقف وفي كل مرة تفكيرا مستقلا.

2- الاتجاه يوجه استجابات الفرد للأشخاص والأشياء والموضوعات بطريقة شبه ثابتة.

3- الاتجاهات تنعكس في سلوك الفرد وفي أقواله وأفعاله وتفاعله مع الآخرين في الجماعات المختلفة في الثقافة التي يعيش فيها.

4- الاتجاهات المعلنة تعبر عن انصياع الفرد لما يسود مجتمعه من معايير وقيم ومعتقدات.

5- الاتجاه يحمل الفرد على أن يحس ويدرك ويفكر بطريقة محددة إزاء موضوعات البيئة الخارجية.

6- الاتجاهات تبلور وتوضح صورة العلاقة بين الفرد وعالمه الاجتماعي فهي تساعده على التكيف مع الحياة الواقعية، كما تساعده على التكيف الاجتماعي وذلك عن طريق قبول الفرد للاتجاهات التي تعتنقها الجماعة فيشاركهم فيها ومن ثم يشعر بالتجانس معهم.

7- تعمل الاتجاهات على إشباع كثير من الدوافع والحاجات النفسية والاجتماعية ومن هذه الحاجات الحاجة إلى التقدير الاجتماعي، والقبول الاجتماعي، والانتماء إلى جماعة معينة، والمشاركة الوجدانية، وهنا يتقبل الفرد قيم الجماعة، ومعاييرها.

8- تعمل الاتجاهات على تقديم المعونة في بلوغ الأهداف في كافة الميادين فإذا كان مدير المدرسة يعبر عن اتجاهاته أما رؤسائه فيجد أن اتجاهاته تتعزز عندما يعبر عنها أمام زملائه، فبذلك تكون اتجاهاته مهمة لبلوغ أهدافه المهنية.

9- تمتلك وظيفة الدفاع عن الذات فالعامل الفقير الأبيض في أمريكا يشعر بأنه ليس أحقر الناس على الأرض عندما يعبر عن عدائه للسود ويدعي بأنهم أحقر منه.

الفرق بين الاتجاه والعاطفة

- **العاطفة:** تمتاز بأنها شخصية وذاتية، فعاطفة الأم نحو أبنائها تختلف عن اتجاه الأم نحو عملها وتقتصر على الجانب الشعوري والوجداني
- **الاتجاه:** أكثر شمولا وعمومية، ويشمل على الجوانب العقلية والمعرفية والإدراكية وسلوكية متعددة.

الاتجاه والرأي العام

- **الاتجاه:** يعني الاستعداد العقلي للاستجابة أو الميل نحو الاقتراب أو الابتعاد عن موضوع ما أي أن الاتجاه يشير إلى ما نحن على استعداد لعمله.
- **الآراء:** تشير إلى ما يعتقد على إنه صواب.

وعلى ذلك فالاتجاهات أكثر عمومية من الآراء التي هي وسيلة التعبير اللفظي عن الاتجاهات.

وقد تستخدم في قياس الآراء سؤال واحد يتضمن كل المعلومات المطلوبة وهذا ما يحدث عند إجراء الانتخابات البرلمانية في البلد حيث يختار الناخب اسم المرشح الذي يرغب في أن يمثله في البرلمان مع قائمة تتضمن مرشحين أو أكثر ولا يطلب من (الناخب) ترتيب المرشحين حسب درجة تفضلية أو يسأل عن مدى شعوره القوي نحو من يرشحه وهذا مثال نموذجي للطرائق المستخدمة في استطلاع رأى الجمهور سواء كانت تتعلق بالسياسة أو ببرامج الراديو أو القضايا التجارية.

الاتجاه والتعصب

التعصب اتجاه سلبي أو ايجابي نحو قضية أو فكرة لا تقوم على أساس منطقي ولم يقم الدليل على صحتها، ويتصف بأنه مشحون بشحنة انفعالية زائدة تجعل التفكير بعيدا عن الموضوعية والمنطق السليم. فتعصب الفرد نحو جماعته يجعله يشعر بالحب نحوها والبغض تجاه كل ما عداها من المجموعات.

العلاقة بين الاتجاهات والقيم

يرى العالم Campell 1963 أن الاتجاهات والقيم حالات مكتسبة من (حالات الدافعية) أي دوافع مكتسبة، وتتدخل معانيها كالقيم والاتجاهات والمعتقدات والإطار المرجعي وغيرها... وعلى الرغم مما يربط الاتجاهات بالقيم مع روابط وعلاقات فإن بينهما اختلافات مميزة تنحصر بالنواحي السبعة التالية:

1- يؤكد روكيتش 1968 أن الاتجاهات والمعتقدات تتعدد حتى تصل إلى آلاف بينما القيم تقل فلا تتجاوز العشرات.. القيم تتخذ تسلسلا هرميا يختلف اتجاهه وأولوياته من شخص إلى لآخر.

2- الثقافة دوما تتصل بالقيم ولا يقال أن للثقافة اتجاهات نفسية وانما لها قيم محددة.

3- القيم هي (النواة) أما الاتجاهات فتتجمع حولها لتوجيه السلوك على مدى طويل نحو هدف له جاذبيته الخاصة.

4- الاتجاهات أكثر عرضة للتغير السريع بينما القيم اشد ثباتا.

5- قد يحصل تعارض بين قيمة معينة واتجاهات متعارضة في شخص واحد رغم علاقات الترابط بينهما فهناك من يساعد الفقراء كالتزام اجتماعي وهذا اتجاه بينما يكون لديه قيم أخرى غير قانعة بجدوى الصدقات فهو يفضل تأسيس معهد لتدريب الفقراء ليكسبوا قوتهم بسواعدهم.

6- أن التعارض بين (أنسقة القيم) تكون عند الفرد اتجاهات متناقضة عن القيم لها تبريراتها عقليا، فالمثقفون يؤيدون الديمقراطية وينادون بالانتخابات ويعاملون زوجاتهم بطريقة معاكسة لا تختلف عن الطريقة التي كان يعاملهن بها الأجداد.

7- كشف العالم روكيتش 1968 عن العلاقة بين القيم والاتجاهات والسلوك في تمييز نوعين من القيم:

أ- القيم الوسيلية: كسعة الأفق، النظافة، العضو، المسؤولية.

ب- القيم الغائية: كالمساواة والسعادة والحرية.

تكوين الاتجاهات

تتكون الاتجاهات نتيجة لاتصال الفرد بالبيئة الطبيعية والاجتماعية المحيطة به وتدور الاتجاهات في بدء نشأتها نحو الأمور المالية، كحب النادي الذي نجد فيه مكانا مريحا أو حب الأسرة أو حب نوع معين من الأطعمة. وتتصف موضوعات الاتجاهات في بادئ الأمر بأن تكون محدودة حيث ينحصر اهتمام الفرد في أفراد من جماعات صغيرة كجماعة الأسرة أو النادي أو رفاق الصف وبعد ذلك تتسع دائرة الاتجاهات وتشتمل على موضوعات مجردة وأمور معنوية وتتكون الاتجاهات النفسية نتيجة لتكامل مجموعة من الخبرات الجزئية التي تدور حول موضوع معين كتكامل هذه الخبرة في وحدة كلية ينتج عنها نوعا من التعميم.

وقد حدد البورت خطوات لتكوين الاتجاه:

1- مرور الفرد بخبرات فردية جزئية تدور حول موضوع الاتجاه.

2- تكامل هذه الخبرات وتنافسها واتحادها في وحدة كلية.

3- تمايز هذه المجموعة من الخبرات وتفردها عن غيرها وظهورها على شكل اتجاه عام.

4- تعميم هذا الاتجاه وتطبيقه على الحالات والمواقف الفردية التي تجابه الفرد والتي تدور حول موضوع الاتجاه.

وهناك طرق كثيرة يكتسب الفرد بواسطتها اتجاهاته. وكل اتجاه يكتسبه الفرد تحدده أمور ثلاثة:

1- تقبل المعايير الاجتماعية بدون نقد ويكون ذلك عن طريق الإيحاء

2- تعميم الخبرات الشخصية

3- الخبرات الانفعالية الشديدة

أما عن العامل الأول وهو الإيحاء، فيعتبر من أكثر تلك العوامل شيوعا ذلك إنه كثيرا ما يقبل الفرد اتجاها مادون أن يكون له اتصال مباشر بالأشياء أو الموضوعات المتصلة بهذا الاتجاه. فالطفل منذ ولادته يقع تحت تأثير الأسرة وهي الجماعة الاجتماعية الأولى التي يعيش الطفل معها ويتفاعل وإياها فعن طريق التقليد يكتسب الطفل كثيرا من العادات والاتجاهات والميول والقيم والمعايير من أفراد الأسرة فيتقبلها وتصبح معاييره أو يرفض البعض الآخر. وتتم عملية التقليد وقبول الآراء بطريقة لاشعورية أو بطريقة شعورية.

مثال: أجريت دراسة لمعرفة الفرق بين اتجاهات الأطفال البيض نحو الزنوج في الأقاليم الشمالية واتجاهات الأطفال البيض نحوهم في الأقاليم الجنوبية ودرست هذه الاتجاهات في مجموعات مختلفة (المناطق الصناعية - الريفية - المقاطعات المكتظة بالزنوج - مدارس الأطفال البيض فقط - مدارس الأطفال البيض والزنوج - أطفال ينتمون إلى أباء ذو نزعة اشتراكية)

وقد دلت النتائج على أن كل البيئات السابقة ماعدا الاشتراكية لها تعصبا نحو الزنوج ويختلف في الدرجة دون النوع أما في البيئات الاشتراكية فإن الأطفال لم يوجد لديهم مظهر من مظاهر التعصب نحو الزنوج. وتدل النتائج على أن الاتجاه أو تكوين رأى ما هو شيء مكتسب أو يتعلم نتيجة الاختلاط بالزنوج أو عدم الاختلاط بهم كما تحدده المعايير الاجتماعية العامة التي يمتصها الأطفال عن آبائهم دون نقد أو تفكير فتصبح جزءا نمطيا من تقاليدهم وحضارتهم يصعب التخلص منها.

أما الوسيلة الثانية فهي تعميم الخبرات فنحن دائما نستعين بخبراتنا الماضية وتعمل على ربطها بحياتنا الحاضرة، فالطفل مثلا يدرب في صغره على عدم الكذب أو عدم اخذ الشيء ليس له أو عدم الاعتماد على جاره في المدرسة في نقل الأعمال المدرسية كالحساب والإنشاء... الخ والطفل ينفذ إرادة والديه في هذه النواحي دون أن تكون لديه أية فكرة عن أسباب ذلك، دون أن يعلم إنه إذا خالف

ذلك عد خائنا وغير أمين. ولكنه عندما يصل إلى درجة من النضج يدرك الفرق بين الأعمال الأخرى التي يوصف فاعلها الخيانة، وحينما يتكون عنده هذا المبدأ أو ذلك المعيار يستطيع أن يعممه في حياته الخاصة والعامة.

أما الوسيلة الثالثة التي يكتسب بها الفرد اتجاهاته فهي عن طريق بعض المواقف ذات الأثر الشديد في نفسه.

المجالات والأغراض التي تستخدم بها مقاييس الاتجاه

1- تستخدم في بحوث استطلاع الرأي العام (الجمهور) حول بعض القضايا السياسية، الاجتماعية، الاقتصادية، وأسئلة عالمية لميول المجتمع وهنا يوجه الاستفتاء الشامل لكافة أرجاء الدولة ويمكن وضع نتائج هذه الاستطلاعات في عدد من الاستخدامات العملية.

2- مقاييس الاتجاه تستخدم في توضيح الاتجاهات الأكثر اتساعا والمعرفة عموما كما هو الحال بالنسبة للفلاحين والأشخاص الذين يعيشون في مناطق واسعة.

3- تستخدم مقاييس الاتجاه في البحوث التجارية لدراسة رأى الجمهور المستهلك لبعض السلع الضرورية، ومعرفة الحاجات الأساسية للمستهلك وردود التي يكونوها عند مراجعة الإنتاج، الخدمات، الإعلانات التجارية، ونتائج المعلومات تستخدم لعدة أغراض كاختيار الإعلان الأكثر تأثيرا بالنسبة للمواد المحددة ولأعداد النموذج الجديد، أو تصميم المنتج الجديد.

4- تستخدم مقاييس الاتجاه في تقدير الاتجاه الكلي للفرد (المؤيد أو المعارض) نحو العمل أو الجماعة.

فعلى سبيل المثال أداة صممت لأغراض (SRA) قائمة المستخدم (الأجير) من (78) فقرة اختيرت كعينة لتمثل الاتجاهات في كل مجال ومتطلبات العمل، العلاقات الشخصية للمستخدم – والمرف على العمل، الدفع، كفاية المستلزمات والآراء المتبادلة والتفاعل مع الجماعة.

5- مقاييس الاتجاهات تستخدم في مجال التربية والتدريب، فمثلا يمكن قياس التغييرات لاتجاهات الطلبة نحو الفن، الاختلافات العرقية، العوامل الحضارية، القضايا الاجتماعية والاقتصادية، وكل ما هو وثيق الصلة بالموضوع، تتبعها عادة إعطاء فصل للدراسة أو برنامج تربوي آخر.

6- ومن أكثر التطبيقات الشائعة لمسوحات الاتجاه نجدها شائعة في بحوث علم النفس الاجتماعي، نظريات الشخصية، والمجالات المتعلقة بها وفي كل كتاب مدرسي نجد به فصول تطبيقية عن الاتجاهات وقياساتها وخاصة كتب علم النفس الاجتماعي.

قياس الاتجاهات

تشير انستازي إلى أن عملية قياس الاتجاهات هو موضوع يثير الجدل والخلاف، فالآراء المعبرة لفظيا يمكن اعتبارها كمؤشرات إلى حقيقة الاتجاهات والتي كثيرا ما أثارت التساؤل في هذا الجزء أن المشكلة معنية أو مهتمة بالعلاقة بين السلوك الصريح اللفظي وغير اللفظي. وبكلمات أخرى هل أن الشخص أفعاله تلائم أقواله أو كلماته عند قياس درجة الاتجاه؟ أن التعارض في التعبير عن الاتجاه سواء بشكل لفظي أو بالسلوك الظاهري تم ملاحظته في العديد من الدراسات. ففي دراسة لاستقصاء اتجاهات الطلبة في الجامعة نحو سلوك الغش في الأوراق الامتحانية وقد كان المقياس المصمم سريا. وقد وجد إنه على الرغم من الاتجاه القوي ضد الغشاش المحترف في الجماعة إلا أن الآراء اختلفت وتنوعت عن الغش وفي هذا الجزء لا يمكن تحديد السلوك اللفظي وغير اللفظي. وهناك مؤشرات أخرى تبين لنا بأنه حتى الملاحظة للسلوك الظاهري هو ليس دائما يمدنا بدقة أو صحة محتوى الاتجاه فمثلا الشخص الذي لديه قوة في المعتقدات الدينية فنراه يلازم الكنيسة دائما ليس بسبب إيمانه الراسخ بالدين وإنما كمعاني للقبول الاجتماعي داخل جماعته. فالعلاقة بين ما يقوله الشخص وما يفعله اعتبرت كحالة خاصة من الصدق.

صدق مقاييس الاتجاه

لقد ذكر ماك تيمان خمس طرائق لإيجاد صدق مقياس الاتجاه:

- ففي الطريقة الأولى يقارن الاتجاه - كما تدل عليه درجاته - مع السلوك الممكن ملاحظته في موقف مطابق ويؤخذ المدى الذي يتفقان فيه كتقدير لصدق المقياس.
- أما الطريقة الثانية لتقدير الصدق مدى تمييز المقياس بين أفراد المجموعات التي تعرف آراؤها، وقد استخدم ثيرستون وشيف هذه الطريقة للحصول على دليل للصدق لمقياسهم للاتجاه نحو الكنيسة ووجدوا إنه يميز بين طلاب علم اللاهوت والطلاب الآخرين، وبين الطلاب ممن لهم انتماءات دينية مختلفة وبين أعضاء نشيطين في الكنيسة وآخرين ممن لهم ميل اقل في شؤون الكنيسة.
- أما الطريقة الثالثة فتتضمن لاستخراج الصدق ارتباط درجات الاختبار بتقديرات الاتجاهات التي يعطيها من لهم معرفة شخصية للأفراد المختبرين. ويوجد هنا بالإضافة إلى عدم الدقة التي يحتمل أن تنتج عن استخدام التقديرات الذاتية من هذا النوع، مصدر آخر للخطأ يرجع إلى حقيقة أن التقديرات جميعها سوف لا تجرى بواسطة الشخص نفسه، ويبدو أن هذه الطريقة في جملتها ترتبط بمصادر كثيرة للخطأ مما يجعلها غير مفيدة كثيرا.
- الطريقة الرابعة المقارنة بمقياس معروف، وقد يبدو أنه لا داعي لأعداد مقياس آخر إذا وجد مقياس له صدق مرض.
- أما الطريقة الأخيرة فهي المقابلة الشخصية للأفراد (المختبرين) لرؤية ما إذا كانوا قادرين على الصمود عند إعادة استجوابهم في آرائهم التي سبق وتعلموها. وتتضمن هذه الطريقة مرة أخرى عنصرا ذاتيا كبيرا. وقد عدل ثيرستون وشيف 1929 في طريقتهم عندما حصلوا على تقديرات ذاتية للاتجاه نحو الكنيسة من أفراد عينتهما فقد حصلا باستخدام خط بياني لمقياس التقديرات على تقديرات ذاتية

قياس الاتجاهات النفسية

الطرق اللفظية لقياس الاتجاهات

تعد الطرق اللفظية لقياس الاتجاهات النفسية من أكثر الأساليب تقدما، نظرا لاعتمادها على الاحتكاك المباشر بالحالة أو الحالات التي يراد قياسها والحصول على إجابات لعدد كبير من الأشخاص في وقت قصير ويمكن تلخيصها في الطرق التالية:

أولا: طريقة الانتخاب Voting

تعتبر هذه الطريقة من الطرق العامة التي يسهل استخدامها وتحليل نتائجها وهي تعتمد على الاستفتاء الذي يتكون من مجموعة أسماء أو موضوعات على صورة مواقف اجتماعية. وعلى الفرد أن يختار أحب هذه الموضوعات إلى نفسه وأهمها لديه أو ابغضها عنده أو غير ذلك من النواحي التي يراد قياسها، ثم بعد ذلك يقوم الباحث بحساب النسبة المئوية للأصوات ثم يرتب الموضوعات ترتيبا يعتمد على القيم العددية لتلك النسب المختلفة وإذا كانت هذه الطريقة تتسم بالسهولة أو السرعة في التطبيق والتحليل والنتائج إلا أنها لا توضح بدقة الفروق القائمة بين موضوعات الاستفتاء وتتأثر كثيرا بالعوامل الخارجية التي لا يشملها أو يتضمنها الاستفتاء.

ثانيا: طريقة الترتيب Rank order

تعتمد هذه الطريقة على أسلوب ترتيب موضوعات الاستفتاء ترتيبا يعتمد في جوهره على نوع الاتجاه المراد قياسه. وبذلك يتكون الاستفتاء من عدد محدود من الموضوعات وتتلخص استجابة الفرد المفحوص في ترتيب هذه الموضوعات بالنسبة للدرجة كلية نحوها أو نفوره منها أو غير ذلك من الاتجاهات النفسية والاجتماعية وفيما يلي مثال لتوضيح ذلك:

طلب من شخص ما ترتيب الألوان المبينة أدناه حسب درجة ميله نحوها وحبه لها بحيث يصبح أولها هو أحب الألوان إليه، وآخرها ابغض الألوان إليه (الأسود - البني - الرمادي - الأصفر - الأحمر - الأخضر - الأزرق - الأبيض) وقد جاءت استجابة الفرد المفحوص على النحو الاتي (الأبيض- الأزرق - الأخضر - الأحمر - الأصفر - الرمادي - البني - الأسود) وبذلك يسفر هذا الاختبار عن تفضيل الشخص المستجوب اللون "الأبيض" على اللون "الأزرق" و"الأزرق" على "الأخضر" و"الأخضر" على اللون "الأحمر" وهكذا حتى ينتهي الترتيب باللون "الأسود" كأبغض لون لهذا الفرد.

ثالثا:طريقة المقارنة المزدوجة Paired comparison

أن خلاصة هذه الطريقة هي أن يفضل الفرد اتجاها على آخر نحو الموضوع الذي يتم قياسه فمثلا إذا أردنا أن نتعرف على اتجاه الفرد من حيث ميله، أو نفوره من حيوانات أو طيور مختلفة فإننا نعرض عليه نوعين من الحيوان ليفضل إحداهما على الأخر. ثم نعرض مرة أخرى عليه حيوانين آخرين ليفضل كذلك واحدا على الآخر. وهكذا وفي هذه الحالة لابد من إعطاء الشخص المستجوب فرصة التفضيل لجميع المقارنات الزوجية الممكنة حيث أن هذه الطريقة تعتمد بالأساس على المقارنات الازدواجية بحيث يتكون كل سؤال من أسئلة الاستفتاء من مقارنة موضوعين ثم تفضيل إحداهما على الأخر بالنسبة لهدف الاتجاه المطلوب قياسه. ويمكن تحليل نتائج هذا القياس بحساب عدد مرات اختيار الفرد وتفضيله لكل موضوع ثم حساب النسبة المئوية لذلك.

رابعا:طريقة التقدير Rating Method

وتعتبر هذه الطريقة من أكثر الطرق استخداما في (قياس الاتجاهات النفسية) وتستخدم في عدة صور منها:

1- مقياس بوجاردس

2- مقياس ثيرستون
3- مقياس ليكرت
4- مقياس كتمان

1- مقياس بوجاردس

لعل أول محاولة لقياس الاتجاهات هي تلك الدراسة التي قام بها بوجاردس للمسافة الاجتماعية Social distance عام 1925. والتي أراد بها التعرف على مدى تقبل الأمريكيين أو نفورهم من أبناء القوميات الأخرى أو على مدى التباعد الاجتماعي بين الأمريكيين من ناحية وأبناء الشعوب الأخرى من ناحية ثانية. في هذه الدراسة نجد أن بوجاردس قد افترض أن العبارات أو الاستجابات السبعة التالية تمثل مسطرة متدرجة للتقبل الاجتماعي، بمعنى أن الطرف الأول من هذا المقياس (العبارة الأولى التي تتحدث عن استجابة تقبل الزواج من احد أفراد هذه القومية) يمثل أقصى درجة من درجات التقبل أو التقارب الاجتماعي. كما افترض أن العبارة السابعة التي تتحدث عن استجابة استبعاد أبناء هؤلاء القومية من الوطن تمثل أقصى درجة من درجات التباعد أو النفور الاجتماعي وان العبارات التي بين الطرفين تمثل درجات متوسطة بين التقبل الاجتماعي والتقارب الاجتماعي، وان بعد أية عبارتين متجاورتين مساو تماما للبعد بين أية عبارتين متجاورتين أخريين.

هذه الوحدات السبعة هي:

1- اقبل أن أتزوج من فرد منهم.
2- اقبل انضمام فرد منهم إلى النادي والذي انتمي إليه ليكون صديقي من بعد ذلك.
3- اقبله جارا لي في المسكن.
4- اقبله واحدا من أبناء مهنتي وفي وطني.
5- اقبله زائرا لوطني
6- اقبل استبعاده من وطني

ثم طالب بوجاردس عينة تتألف من 1725 أمريكيا أن يحددوا اتجاهاتهم نحو أبناء عدد من القوميات باستخدام هذا المقياس فخرج بعدد من النتائج منها على سبيل المثال: كانت نتائج تطبيق هذا المقياس بالنسبة لأربعة من الشعوب كما موضحة في الجدول أدناه:

الشعوب / درجات المقياس	1	2	3	4	5	6	7
الإنكليز	93,7	96,7	97,3	95,4	95,9	1,7	–
السويد	45,3	62,1	75,6	78,0	86,3	5,4	1,0
البولنديون	11,0	11,6	28,3	44,3	58,3	19,7	4,7
الكوريون	1,1	6,8	13,0	2,4	23,7	47,1	19,1

ويلاحظ أن هذه النتائج لا تمثل مقياسا واحدا بل مجموعة مقاييس فهو مقياس لكل شعب من الشعوب المختلفة وحداته الاستجابات السبعة. ويتضح أن تدرج وحدات هذا المقياس (ليست متساوية البعد) وخاصة بعد الوحدة الرابعة فهو مقياس تجمعي أي الذي يوافق على الوحدة الأولى يوافق على الثانية والرابعة والخامسة، ولذا فإن النسبة في تزايد مستمر حتى الوحدة الخامسة. بينما الذين يوافقون على السادسة أو السابعة هم فئة أخرى لا علاقة لها بالذين يوافقون على الوحدات الأولى.

وهناك طرق عديدة قد اقترحت لإعطاء درجة لكل شعب من هذه الشعوب بناء على الاستجابات التي يعطيها الأفراد الذين يجرى عليهم المقياس ومنها الطريقة الموضحة في الجدول التالي.

جدول يوضح طريقة إحصائية لمقارنة المجموعات في طريقة البعد الاجتماعي

الوحدة/ الشعوب رقم	الانكليز	السويد	البولنديون	الكوريون
1	1 ×93,7	1×45,3	1×0,11	1×1,1
2	2×96,7	2×62,1	2×11,6	2×10,8
3	3×97,3	3×75,6	3×28,3	3×11,8
4	4×95,4	4×78,0	4×44,3	4×20,1
5	5×95,9	5×86,3	5×58,3	5×27,5
المجموع	1440,1	1339,8	587,8	276,0

وتتلخص الطريقة في إعطاء أوزان للنسب المئوية لاستجابات وحدات المقياس أي بضرب كل نسبة في رقم الوحدة:

ومن هذا نستطيع أن نقول أن ترتيب تقبل العينة المختبرة من الأمريكيين للشعوب وهي كما يأتي: الانكليز- السويد - البولانديون - الكوريون.

إن الخاصية الإحصائية للمقياس الصحيح وهي التدرج تتوفر في هذا المقياس والذي يدل على ذلك تزايد النسب في وحداته بالنسبة للشعوب المختلفة. ويمكن التأكد من ذلك بتطبيق طريقة تحليل التباين على النسب المختلفة في الوحدات السبعة.

أما أهم الانتقادات التي وجهت إلى هذه الطريقة:

1- أن نتائج المقياس تمثل نتائج عدة مقاييس فهو مجموعة مقاييس يمثل كل واحد منها مقياسا لكل شعب من الشعوب ويكون من سبع وحدات.

2- أن وحدات هذا المقياس لا تتدرج بشكل متساو.

3- أن من يوافق على الوحدة الأولى في المقياس يوافق عادة على الوحدات الثانية والثالثة والرابعة والخامسة.

4- لا تشتمل وحداته على مواقف تعكس اتجاهات بعض الأفراد المتطرفين تطرفا زائدا نحو الأجناس أو الشعوب المختلفة.

2- مقياس ثيرستون

وضع ثيرستون عددا من المقاييس لقياس الاتجاهات نحو موضوعات متعددة من ذلك التفرقة العنصرية، الحرية، الحرب، الكنيسة، الزنوج، ا لوطنية، وغير ذلك. وتسمى هذه الطريقة باسم Method of Equal- Appearing intervals طريقة الفقرات المتساوية البعد. ولقد استخدمها Thurstone أول استخدامها في المقياس الذي وضعه مع chave لقياس الاتجاه نحو الكنيسة عام 1929–1931. حيث كان يسعى إلى بناء مقياس ذي بنود منتظمة وموزعة على متصل متساوي المسافات وذلك باستخدام وسيطات Medians هذه البنود التي يتم حسابها نتيجة لتحكيمها من قبل محكمين أكفاء ذوي خبرة ودراية عاليتين. وقد اقترح ثيرستون الخطوات المنهجية المحددة لتصميم هذا المقياس وعلى الوجه الآتي:

1- يطلب عدد كبير من الأفراد تقديم بعض الجمل التي تصف اتجاهاتهم نحو مؤسسة اجتماعية معينة (أو موضوع معين) مثل المؤسسة التعليمية أو الكنيسة أو الحرب أو التفرقة العنصرية ثم تراجع هذه الجمل وتبلور من خلال الكتابات التي يحفل بها التراث عن الموضوع، وتعاد صياغتها بصورة مبسطة ومختصرة لتصبح في شكل جمل قصيرة واضحة المعنى.

2- تقدم هذه المجموعة من الجمل والتي عددها (130) جملة إلى عدد من المحكمين بلغ عددهم (300) محكم الذين كان يطلب منهم تصنيفا من حيث تدرجها في التعبير عن الاتجاه نحو المؤسسة المعنية في (11) فئة من A--------K بحيث تكون الفئة (A) هي التي تتضمن الجمل الأكثر إيجابية وتأييدا وقبولا لهذه المؤسسة وتليها الفئة (B) إلى الفئة (F) التي تمثل موقفا محايدا من هذه المؤسسة التي تتدرج الجمل في الفئات من G----K من الاتجاه السلبي، والرافض إلى أقصى درجات الرفض التي تعبر عنها الجمل المصنفة في الفئة (K). ويلاحظ أن مهمة المحكمين

هنا ليست التعبير عن آرائهم أو اتجاهاتهم نحو هذه المؤسسة أو الموضوع أو اختيار الجمل التي تعبر عن اتجاهاتهم بل أن يحددوا فقط مدى تعبير هذه الجمل عن اتجاهات معينة دون اعتبار لموقفهم الشخصي.

3- بعد هذه الخطوة يقوم الباحث بتحويل التصنيف من حروف أبجدية إلى درجات هي إحدى عشر درجة أيضاً وبذلك تمنح الجملة التي أعطيت التقدير K إحدى عشر درجة أما الجملة التي أعطاها الحكم التقدير A فتمنح الدرجة واحد صحيح وهكذا بالنسبة للجملة المحايدة فتمنح الدرجة (6) بعد عملية تحويل الحروف إلى أرقام. يقوم الباحث بتحليل الدرجات المعطاة لكل جملة على حدة ثم يستخرج متوسط هذه الدرجات أو وسيطها وتصبح هذه الدرجة هي القيمة التي أعطاها جميع الحكام لهذه الجملة وقد أطلق ثيرستون اصطلاح Scale-Value.

4- ولمعرفة مدى اتفاق الحكام حول معنى الجملة فإن الباحث يقوم بقياس الفروق الفردية في الدرجات المعطاة لها وذلك عن طريق إيجاد الانحراف المعياري لكل جملة والانحراف المعياري هو مقياس درجة تشتت أو انتشار الدرجات.

5- الجمل التي يختلف حولها الحكام اختلافا كبيرا أي الجمل التي بها انحراف معياري كبير، هذه الجمل تحذف من الصورة النهائية للمقياس لان معناها ليس واحدا بالنسبة لجميع الأفراد فالفقراء يختلفون فيما بينهم في درجة تعبيرها، عن الاتجاه بعبارة أخرى تعتبر هذه الجمل غامضة وغير واضحة.

6- يختار الباحث عددا من الجمل الواضحة والتي تنتشر انتشارا متساويا وعلى المقياس من الطرف المؤيد إلى الطرف المعارض.

7- وفي الغالب المقياس النهائي يتكون من حوالي 20–25 جملة، يقدم الباحث المقياس في صورته النهائية للمجموعة المراد قياس اتجاهها ثم يؤخذ وسيط جميع الجمل التي أجاب عنها المفحوص ويعبر هذا الوسيط عن درجته. أما مقياس ثيرستون فلم يتمكن من استبقاء سوى (45) جملة متساوية البعد تقريبا واليك أمثلة من هذه الجمل والرتب التي أعطيت لها في المقياس:

	الجمل	الأوزان
1.	اعتقد أن الكنيسة أعظم معهد أمريكي في الوقت الحاضر	0,2
2.	اعتقد أن الانتماء إلى الكنيسة يعتبر تقريبا من أساسيات الحياة الناجحة في الوقت الحاضر.	1,5
3.	أرى أن الخدمات التي تقدمها الكنيسة تبعث على الاستقرار والأمل.	2,2
4.	أحب كنيستي لان فيها يشيع جو المحبة	3,3
5.	اصدق ما تقدمه الكنيسة من تعليم ولكن مع بعض التحفظ العقلي	4,5
6.	أشعر في بعض الأحيان بأن الكنيسة والدين ضروريان، بينما أشك في ذلك أحيانا أخرى.	5,6
7.	أظن أن الكنيسة معطلة للدين لأنها لاتزال تعتمد على الخرافات السحرية والأساطير	9,6

أما عن الطريقة التي يستطيع الباحث إتباعها لتحديد مدى انتشار الدرجات المعطاة لكل قضية من قضايا المقياس فإن هناك طريقة عملية لإيجاد مقياس معقول لانتشار الدرجات وذلك باستخراج المدى الربيعي inter Quartile range ويمكن تلخيص العملية الحسابية في الخطوات الآتية:

1- عمل توزيع تكراري للدرجات المعطاة لكل جملة.

2- إيجاد عدد الحكام الذين أعطوا للمجموعة درجة معينة.

3- تحويل هذا العدد إلى نسبة بقسمته على عدد الحكام وبذلك تصبح نسبة Proportion – إيجاد توزيع تكراري تجمعي لهذه النسبة Cumulative Frequencies

4- عمل رسم بياني حيث توضح قاعدة الرسم الدرجات المعطاة للجملة على المقياس ذي الـ (11) نقطة والمحور الرأسي يوضح النسب التجميعية الناتجة من تقديرات الحكام.

5- بإسقاط عمود على المحور الأفقي من عند نقطة 50٪ من النسب التجمعية سوف يعطى هذا قيمة الوسيط Median وسوف تعتبر هذه القيمة هي القيمة المعطاة للقضية بواسطة جميع الحكام.

6- بإسقاط عمودين عند نقطتي 0,25 و75.0 من عند النسب التجميعية على المحور الأفقي أيضاً سوف يعطيان الرباعي الأدنى والأعلى على التوالي والمسافة بين هاتين القيمتين على المحور الأفقي تعطي قيمة المدى الربيعي الذي هو مقياس للانحرافات وبالتالي فإنه مقياس درجة وضوح أو غموض الجملة.

ومعنى هذه العملية أن 50٪ من الحكام قد اتفقوا في إعطاء درجة معينة لقضية ما ولذلك تؤخذ هذه القيمة للتعبير عن الجملة، وكذلك تدلنا قيمة المدى الربيعي على مدى اختلاف الحكام فكلما زادت هذه القيمة كلما دل ذلك على أن الحكام لم يتفقوا على معنى الجملة ومدلولها بالنسبة للاتجاه.

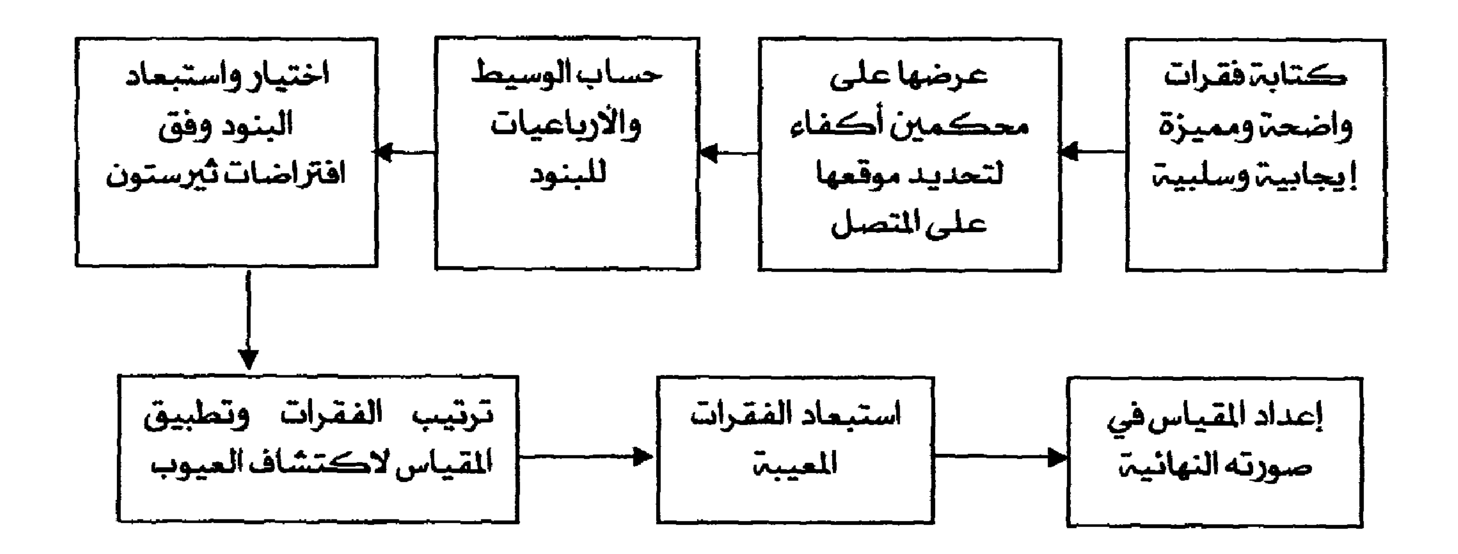
العدد اللازم من المحكمين

لا يوجد بين علماء النفس الذين طبقوا هذه الطريقة اتفاق على عدد الحكام المناسبين وان كان ثيرستون قد استخدم (300) حكما في تصميم مقياس الاتجاه نحو الكنيسة ولكن كثيرا من الأبحاث استخدمت أعداداً اقل من ذلك بكثير وحصلوا على تقديرات صادقة ولقد وصل هذا العدد إلى (30) حكما فقط. Thurstone and chave استخدما (300) محكم في الحصول على قيم المقياس ل(130) جملة عند بناء مقياس الاتجاه نحو الكنيسة. وأشارت البحوث إلى إنه بالإمكان الحصول على ثبات قيم المقياس باستخدام مجموعة صغيرة جدا من المحكمين. فقد أشار Kenny and Edwards أن معامل الارتباط بين قيم المقياس كان 0,95 لمقياس متكون من (129) جملة حصل عليها من تقدير (72) محكم لمقياس الاتجاه نحو الكنيسة الذي بناه ثيرستون وشيف. Undrock 1934 حصل على (279) جملة من مجموعتين للمحكمين في كل

مجموعة (50) محكم، وكان معامل الارتباط بين قيم المقياس والذي حصل من كل مجموعة على التوالي هو 0,99 .

هذه الدراسات تبين لنا بأن عدد قليل من المحكمين بالإمكان استخدامهم للحصول على ثبات قيم المقياس للجمل المستخدمة بهذه الطريقة أن تقليل عدد المحكمين من 300الى 50 والى 30 يقلل لنا الوقت والجهد في الحصول على المحكمين.

والمخطط التالي يوضح طريقة بناء مقياس الاتجاهات على وفق طريق ثيرستون:



الانتقادات التي وجهت لهذه الطريقة:

1- تقوم هذه الطريقة على أساس افتراض أن المسافات بين الجمل متساوية ولكن في الواقع لا تمدنا هذه الطريقة بأي دليل على صحة هذا الغرض.

2- أن الحكم لا تتاح له الفرصة لتغيير رأيه في أثناء القيام بتصنيف الجمل على المجاميع الإحدى عشرة، أو القول بأنه يضطر لوضع جملة مختلفة في (مجموعة واحدة) ولكن يستطيع الباحث التغلب على هذه الصعوبة عن طريق تكليف الحكام بقراءة الجمل أولاً وبعد اخذ فكرة عنها جميعا يبدأون في عملية التقدير (أي تصنيف مقارن للقضايا أو وضع هذه القضايا في رتب معينة).

3- أن غالبية الجمل تميل إلى التجمع حول الطرفين مع بقاء المنطقة المتوسطة من

المقياس خالية وتميل أنواع معينة من القضايا إلى التمركز في المنطقة المحايدة مثل القضايا الغامضة والجمل الخارجة عن الموضوع والجمل المعبرة عن التذبذب وعدم الاستقرار على رأي معين، وان تساوي الدرجة المعطاة لجملة ما يعني أنها متساوية في مضمونها مع جملة أخرى تحمل نفس الدرجة فقد يعبر كل منها عن بعد معين من أبعاد الاتجاه.

ومن مميزات هذه الطريقة:

أن بعض جمل المقياس تكون ايجابية بينما يكون البعض الآخر سلبيا بمعنى أن تكون نصف الجمل مؤيدة والنصف الآخر معارضة وذلك للتغلب على تأثير نمط الاستجابة حيث يميل بعض الأفراد إلى إعطاء استجابة على نمط واحد.

3- طريقة ليكرت (التقديرات المجمعة) Summated Ratings

ابتكر ليكرت طريقته لقياس الاتجاهات في عام 1932 وانتشرت لقياس الاتجاهات نحو مختلف الموضوعات كالزنوج، التقدمية، المرأة، العمل... الخ وقد تميزت طريقة ليكرت على طريقة ثيرستون بما يلي:

1- تتيح لنا طريقة ليكرت اختيار عدد اكبر من العبارات التي ترتبط ارتباطا عاليا مع الاختبار ككل مع أن الحكام قد يختلفون على مدى قيمتها من ظاهر محتواها في قياس الاتجاه موضع الاعتبار. وهذا يتيح لنا تناول جوانب عديدة للاتجاه لا يشملها مقياس ثيرستون.

2- لا تحتاج طريقة ليكرت إلى الحكام ولا إلى اتفاقهم.

3- يمكن لمقياس ليكرت أن يقيس درجات من الاتجاه بالنسبة لكل عبارة. حيث توجد خمس درجات تتراوح بين موافق جدا إلى غير موافق إطلاقا. أما في طريقة ثيرستون فهو أما يوافق الشخص على العبارات أو يتركها وليس هناك مواقف أو درجات في هذه الموافقة أو عدمها.

4- يزودنا مقياس ليكرت بمعلومات أكمل عن المفحوص لأنه يستجيب لكل عبارة

أما اختبار ثيرستون فإن المفحوص حر في أن يترك العبارات التي يعترض عليها.

أما طريقة تطبيق هذه الطريقة تتلخص في الخطوات التالية:

أ- تجمع العبارات من المصادر المختلفة كالجرائد والمجلات والكتب والدراسات السابقة عن المشكلة وتوضع في استمارة خاصة. ولا مانع هنا من استشارة آراء المختصين عن هذه الجمل وما يفهم منها ومدى شمولها على الجوانب المتعددة للمشكلة. وبالرغم من أن هذا المقياس لا يحتاج إلى عدد كبير من الوحدات وقد يكفي (15) وحدة للقياس إلا أن من الأفضل أن يبدأ الباحث بجمع عدد كبير من العبارات.

ب- تخلط هذه العبارات خلطا عشوائيا فبعد تصنيفها تصنيفا متزنا بين ميادين المشكلة المختلفة حتى لا يطغي ميدان من ميادين المشكلة على باقي الميادين في تكوين المقياس تفصل الوحدات المتعلقة بميدان واحد عن بعضها حتى لا يؤثر تجميعها على التزام المختبر لنوع واحد من الاستجابات لها جميعا.

وينبغي أن تتوفر في مفردات المقياس الخصائص التالية كما حددها ثيرستون وشيف وليكرت واخرون.

صياغة فقرات مقاييس الاتجاهات:

هناك بعض الشروط التي يجب أخذها في الاعتبار عند صياغة مفردات الاختبار أو أسئلته أو الجمل التي يتكون منها الاختبار وكما حددها ثيرستون:

1- يجب أن تصاغ الجمل في صيغة الحاضر.
2- تجنب الجمل التي تعبر عن الحقيقة، أو العبارات التي تفسر على أنها حقيقة.
3- تجنب الجمل التي بالإمكان تفسيرها في أكثر من طريقة واحدة.
4- تجنب الجمل التي ليس لها علاقة بالموضوع النفسي.
5- ينبغي عدم استخدام الجمل التي يحتمل أن يوافق جميع أفراد العينة عليها أو التي يحتمل إلا يوافق عليها الجميع.

6- اختار الجمل التي تعتقد أنها تغطي المجال كله في المقياس الذي تريد قياس أثره.
7- أن تكون لغة العبارات بسيطة، وواضحة ومباشرة.
8- الجمل ينبغي أن تكون قصيرة، ولا تتجاوز أل(20) كلمة.
9- كل جملة ينبغي أن تعبر عن فكرة واحدة.
10- ينبغي تجنب استخدام الشوامل مثل كل، دائما، مطلقا، لا احد، أبدا، وينبغي تجنبه
11- عند كتابة الجمل استخدم بعناية ولطف الكلمات مثل، فقط، تماما، تقريبا، ... الخ
12- استخدم الجمل البسيطة بدلا من الجمل المركبة والمعقدة.
13- تجنب استخدام الكلمات غير الواضحة وخاصة في حالة إعطائنا القياس الكامل.
14- تجنب استخدام النفي المزدوج.

ج- تتبع العبارة عادة بخمسة احتمالات للاستجابة بين الموافقة جدا إلى غير الموافقة جدا وإذا استجاب الفرد إلى أي من هذه الوحدات فإنه يعطى الدرجة المناظرة لها وبذلك يمكن حساب درجة الفرد على مقياس الاتجاه بمجرد الجمع البسيط لدرجاته على عبارات هذا المقياس.

ويقوم الباحث بعد ذلك بتحليل فقرات المقياس item analysis وذلك لاستخراج قوة تمييز كل فقرة، واختيار الفقرات المميزة للمقياس، ويقصد بقوة تمييز الفقرة، مدى قدرة الفقرة على التمييز بين المجموعة العليا والمجموعة الدنيا من الأفراد بالنسبة للسمة التي تقيسها الفقرة وقد يستخدم الاختبار التالي (T-Test) لتحليل كل فقرة من فقرات المقياس وذلك للمقارنة بين متوسط درجات المجموعتين العليا والدنيا ولكل فقرة من فقرات المقياس ويمكن أخذ 25٪ بالنسبة للمجموعة العليا و25٪ بالنسبة للمجموعة الدنيا الحاصلة على الدرجات القليلة. في حالة عدم تساوي حجم العينة فيطبق القانون التالي:

$$T = \frac{\xi_L - \xi_{**}}{\sqrt{S^2_{**} / N_{**} + S^2_L / N_L}}$$

حيث إن:

ξ_{**} = متوسط الدرجات على الفقرة للمجموعة العليا

ξ_L = متوسط الدرجات على الفقرة نفسها للمجموعة الدنيا

S^2_{**} = التباين لاستجابات المجموعة العليا على الفقرة

S^2_L = التباين لتوزيع الاستجابات للمجموعة الدنيا على الفقرة

N_{**} = حجم العينة في المجموعة العليا

N_L = حجم العينة في المجموعة الدنيا

في حالة تساوي حجم العينة في المجموعتين العليا والدنيا أي $N_{**} = N_L = N$ واخترنا نفس النسبة فالقانون يكون:

$$T = \frac{\xi_{**} - \xi_L}{\sqrt{sum(X_{**} - \xi_{**})^2 (X_L - \xi_L)^2 / n(n-1)}}$$

وتعبر قيمة (t) أي النسبة الحرجة عن مدى اختلاف العبارة عند كل من المجموعة العليا والدنيا. وعلى الوجه التقريب إذا بلغت قيمة (t) مقدار 1,75 فإن ذلك يدل على أن متوسط الاستجابة لهذه العبارة يختلف اختلاف جوهريا عند المجموعة العليا أو المجموعة الدنيا. أي هناك فرق ذا دلالة إحصائية بين متوسط استجابات المجموعة العليا والمجموعة الدنيا.

وفي طريقة تصميم مقياس ليكرت يكون المطلوب إعداد نحو 20-25 عبارة التي تختلف عند المجموعة العليا منها عند المجموعة الدنيا ويمكن اختبار هذه العبارات عن طريق قيمة (t) وذلك بعد إيجاد قيمة (t) لكن عبارة وتوضع العبارات في ترتيب تنازلي أو تصاعدي طبقا لهذه القيم ثم نأخذ العشرين عبارة ذات اكبر قيمة تائية وتحذف ما عداها. ومعنى ذلك أننا أخذنا العبارات التي يختلف حولها المجموعة

العليا والمجموعة الدنيا، كلما زادت قيمة (t) كلما زاد هذا الفرق وكلما دل ذلك على أن العبارة مميزة ومعبرة فعلا عن الاتجاه.

ويمكن استخدام أي طريقة من طرق تحليل الفقرات مثل معامل الارتباط بين العبارة وبقية الاختبار (أي الدرجة الكلية للاختبار) ويمكن استخدام طرق أسهل من طرق تحليل الفقرات مثل إيجاد الفرق بين متوسطي المجموعة العليا والمجموعة الدنيا.

ولقد وجد كل من Murphy and Likert حين قارنا بين ترتيب (15) عبارة عن طريق الفرق بين متوسط المجموعة المنخفضة والعالية وبين هذا الترتيب الموضوع على أساس معامل الارتباط بين كل عبارة ببقية الاختبار ووجد أن هناك تشابها في هذا الترتيب. والحكمة من استخدام نصف العبارات مؤيدة والنصف الآخر معارضا هو تقليل تأثير نمط الاستجابة ومعنى نمط الاستجابة هو أن هناك ميلا لدى بعض الناس لان يصدروا نمطا معينا من الاستجابة إزاء جميع العبارات فقد يميل الفرد إلى إعطاء الاستجابة المؤيدة باستمرار أو يميل إلى العكس أي إعطاء استجابة معارضة.

أما ثبات الدرجات على المقياس بالإمكان الحصول عليها بواسطة العلاقة أو الارتباط بين عدد الجمل الفردية والجمل الزوجية، أو بطريقة ألفا كرونباك للفقرات التي تم اختيارها فعندما نحصل على قيمة مرتفعة لهذا المعامل فإننا يمكن أن نبقي على هذه الفقرات كصورة نهائية للمقياس. معامل الثبات الخاص بطريقة ليكرت هو أعلى من 0,85 حتى في حالة استخدام (20) فقرة بالمقياس.

وبالإمكان استخراج متوسط درجة الاتجاه للاختبار من أفراد العينة بواسطة المعادلة الآتية:

$$\xi = \frac{\text{sum } X}{N}$$

حيث أن:

ع: الوسط الحسابي

$\sum X$: مجموع الدرجات لكل أفراد العينة على مقياس الاتجاه.

N : عدد أفراد العينة في المجموعات

أسباب تفضيل طريقة ليكرت على الطريقة ثيرستون:

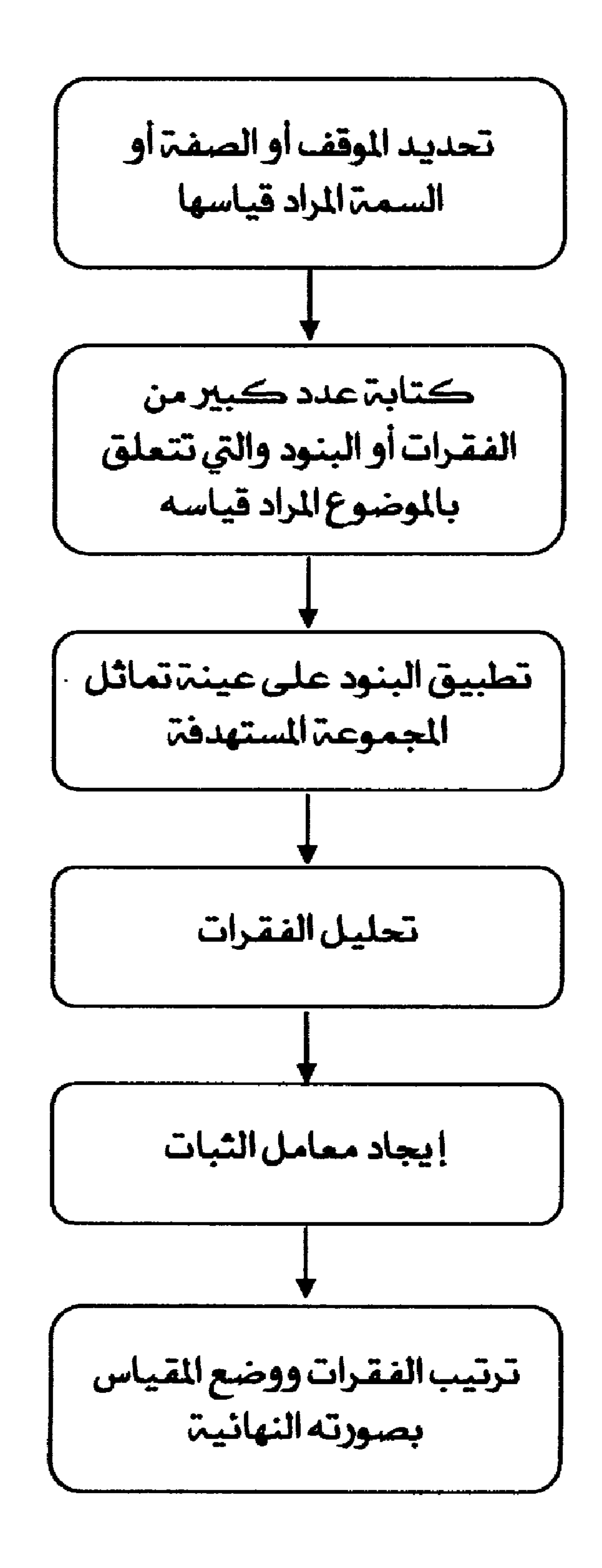
1- أنها أسهل وابسط في تزويدنا بتطور مقياس الاتجاه.

2- أنها اقل جهدا من طريقة ثيرستون.

3- أن الوقت المتطلب في بناء المقياس وفق طريقة ثيرستون هو ضعف الوقت المستخدم في بناء طريقة ليكرت.

4- تستخدم مجموعة كبيرة من الحكام لغرض الحصول على قيم المقياس.

ويمكن توضيح طريقة ليكرت بالمخطط الآتي:


تحديد الموقف أو الصفة أو
السمة المراد قياسها
كتابة عدد كبير من
الفقرات أو البنود والتي تتعلق
بالموضوع المراد قياسه
تطبيق البنود على عينة تماثل
المجموعة المستهدفة
تحليل الفقرات
إيجاد معامل الثبات
ترتيب الفقرات ووضع المقياس
بصورته النهائية

4- طريقة كتمان Guttman التدرج التجمعي Cumulative echnique

حاول كتمان 1947-1950 انشاء مقياس تجمعي متدرج يحقق فيه شرطا هاما هو الحصول على مقياس يقيس صفة أو اتجاه من بعد واحد Unidimenstional ذلك لان كتمان يعتبر الميدان خاضعا للقياس المتدرج التجمعي إذا أمكن ترتيب الاستجابات بطريقة معينة بحيث تجعل من يجيب على إحدى الوحدات بالقبول أعلى مرتبة من الذي يجيب عنها بالرفض. وبذلك يتسنى معرفة نمط إجابته عليها Socilogram analysis.

والمثال التالي يوضح المقصود من هذا المقياس:

1- ينبغي أن تكمل البنت تعليمها حتى بعد الجامعة نعم – لا
2- نهاية المستوى الثانوي لا يعتبر كافيا لتعليم البنت نعم – لا
3- نهاية المستوى الإعدادي لا يعتبر كافيا لتعليم البنت نعم – لا
4- نهاية المستوى الابتدائي لا يعتبر كافيا لتعليم البنت نعم – لا
5- ينبغي أن يزيد تعليم البنت عن مجرد القراءة والكتابة نعم – لا

رتبة المختبر	الإجابة بنعم للجملة رقم					الإجابة بلا للجملة رقم				
	1	2	3	4	5	1	2	3	4	5
1	×	×	×	×	×	–	–	–	–	–
2	–	×	×	×	×	×	–	–	–	–
3	–	–	×	×	×	×	×	–	–	–
4	–	–	–	×	×	×	×	×	–	–
5	–	–	–	–	×	×	×	×	×	–

ففي هذا النموذج نجد إنه إذا أجاب المختبر عن إحدى الوحدات بالإيجاب فلابد أن يجيب بالإيجاب على الوحدات التي بعدها. ومن المتبع عادة زيادة خطوات

مثل هذا المقياس المتدرج بوضع الاستجابة لكل جملة في المقياس على درجات كالدرجات المتبعة في مقياس ليكرت بدلا من نوعين لا، نعم مثل: اؤيد بشدة، أؤيد، لا رأي لي، أعارض، أعارض بشدة. ويتحدد ترتيب المختبر على أساس هذه الدرجات كذلك.

وحدد كتمان الخطوات الآتية في بناء هذه المقاييس:

1- تحديد موضوع الوقف أو السمة أو الاتجاه المراد قياسه ومن ثم كتابة عدد من الفقرات أو البنود التي يجب أن يتوفر فيها شروط معينة بحيث ترتب هذه البنود على متصل من الأقل إيجابية إلى الأكثر إيجابية، وأن تقتضي الموافقة على فقرة ما الموافقة على جميع الفقرات الأقل إيجابية منها.

2- تنظيم وكتابة هذه البنود وإعدادها في صورة قابلة للتطبيق.

3- تطبيق الأداة على عينة يشترط فيها كتمان أن تكون كبيرة نوعا ما بحيث لا ينقص عدد أفرادها عن خمسة أضعاف عدد البنود.

4- تحليل البيانات المتحصل عليها حسب ما يتطلبه كتمان من حساب معامل الإنتاج Coefficient of Reproducibility أو ما يسمى أحيانا معامل إعادة تكوين نمط الاستجابات وكذلك حساب معامل القياسية Coefficient of scalability. وتتضح أهمية حساب هذه المعاملات في بيان مدى اتساق الفقرات مع شروط ومواصفات كتمان. ويمكن استخدام المعادلات التالية لحساب معامل الإنتاج ومعامل القياسية:

- معامل الإنتاج = 1- (عدد الأخطاء ÷ عدد الاستجابات الكلي)
- معامل القياسية = (نسبة التحسن ÷ نسبة التحسن المحتملة)

حيث أن:

نسبة التحسن = القيمة الحقيقية لمعامل الإنتاج – أقل قيمة لمعامل الإنتاج

نسبة التحسن المحتملة = 100 – أقل قيمة لمعامل الإنتاج

ويقترح كتمان إلا يقل معامل الإنتاج عن 90٪ أي أن نسبة الخطأ يجب إلا تتعدى 10٪ بينما يؤكد أن معامل القياسية يجب أن لا يقل عن 60٪ وعند تحقيق هاتين القيمتين فإن المقياس يكون جيدا ومناسبا لأسلوب كتمان.

خامسا: طريقة تمايز معاني المفاهيم The semantic Differential

وضع هذه الطريقة شارلس اوزكود Osgood عام 1957 لدراسة الإدراك والاتجاهات وقياسها بصورة كمية عن طريق طرح معلومات محددة وكمية عن معاني بعض المفاهيم الاجتماعية المراد قياس اتجاهها عند جماعة معينة، وذلك عن طريق وضع درجات للمقياس تفعل بين مفهوم وآخر مثل: الصفة الموجبة (حسن) تمثل الدرجة (7) والصفة السالبة (رديء) تمثل متوسط الصفة.

(1)–2–3–(4)–5–6–(7)

حسن متوسط رديء

إن هذا الأسلوب يمكننا من الحصول على تقديرات للمفاهيم من خلال سلسلة من الموازين للصفات المتضادة ثنائية القطب، وهذا الأسلوب لا يعد أداة قياس معينة أو اختبار، إنما هو أسلوب قياس مرن يستخدم لقياس الاتجاهات وغيرها من الجوانب الوجدانية.

ولبناء هذا المقياس يمكن اتباع الخطوات الآتية:

1- تحديد المفهوم أو المفاهيم المراد تقديرها بأسلوب يعكس ترابطها وعلاقتها بالسمة التي يراد قياسها.

2- اختيار موازين ثنائية القطب مناسبة أو ملائمة لمستوى تفكير الأفراد المستجيبين وممثلة للمفاهيم التي يراد قياسها.

3- تصميم صفحات لتدوين الاستجابات بحيث يظهر كل مفهوم في أعلى صفحة مستقلة متبوعاً بموازين ثنائية ومتبادلة أقطابها عشوائياً، يفصل بينها عدد ثابت من النقاط (7 أو 9) نقاط.

التعلم

سريع – – – – – – – – – – – بطيء

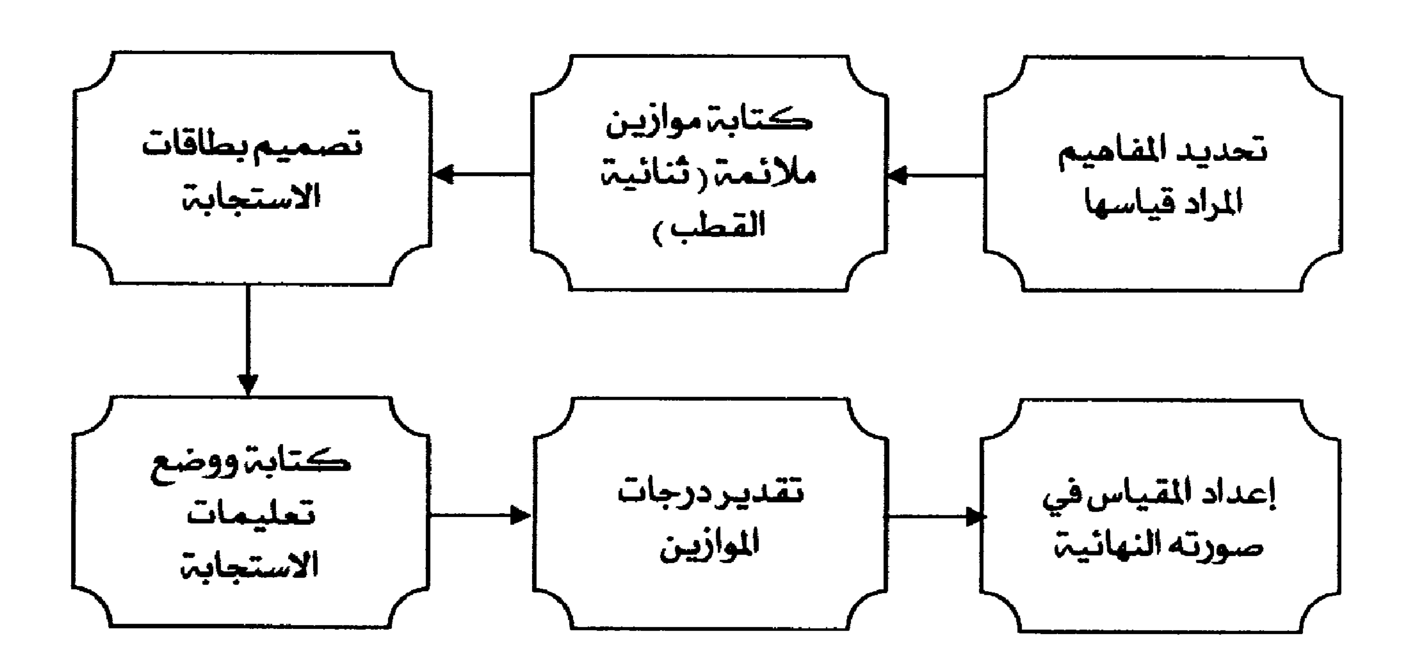
رديء – – – – – – – – – – – جيد

مفيد – – – – – – – – – – – غير مفيد

غير سار – – – – – – – – – – – سار

4- كتابة تعليمات الاستجابة على صفحة الغلاف بحيث توضح المطلوب من المستجيب والكيفية التي يتم بها تدوين الاستجابة.

5- تقدير درجات الموازين بقيم تتراوح مابين (1) إلى (7) أو (1) إلى (9) حسب نقاط الميزان بحيث تقع القيمة الصغرى قريبا من الصفة الممثلة للقطب السالب.



يعتبر هذا الأسلوب أداة جيدة لقياس الجانب الوجداني والمشاعر الإيجابية أو السلبية للناس نحو موضوع معين وبالتالي فهو يصلح لقياس ردود الفعل وليس لقياس الآراء، ويعاب عليه كونه يعطي معلومات عامة ولا يعطي معلومات حول المصدر نفسه وعليه فإن الاعتماد عليه في تقييم البرامج محدود، كما ان من أهم عيوبه صعوبة تقديم مثل هذا النوع من المقاييس للأطفال باعتبارهم غير قادرين على الاستجابة للأسئلة الترتيبية التي تستخدم خطوطاً أو نقاط لوضع الاستجابة.

ويرى علام 2000 أن المقياس صالح لقياس ردود الفعل نحو كثير من المواضيع والأشياء، ويؤكد على ضرورة العناية باختيار الموازين التقييمية حتى لا يكون هناك أخطاء في التقدير، حيث يجب التأكد من أن الصفتين تعتبران طرفي نقيض على متصل متدرج.

هناك تساؤلات حول قدرة هذا الأسلوب الفعلية في قياس المعنى حيث وجد أن بعض المفاهيم تنال تقديرات متماثلة من معظم الأفراد على الرغم من كونها غير متماثلة تماما، ويعود السبب في ذلك إلى اختلاف فهم مضمون الكلمات عند الأفراد تبعاً لتباين خلفياتهم الثقافية والفلسفية.

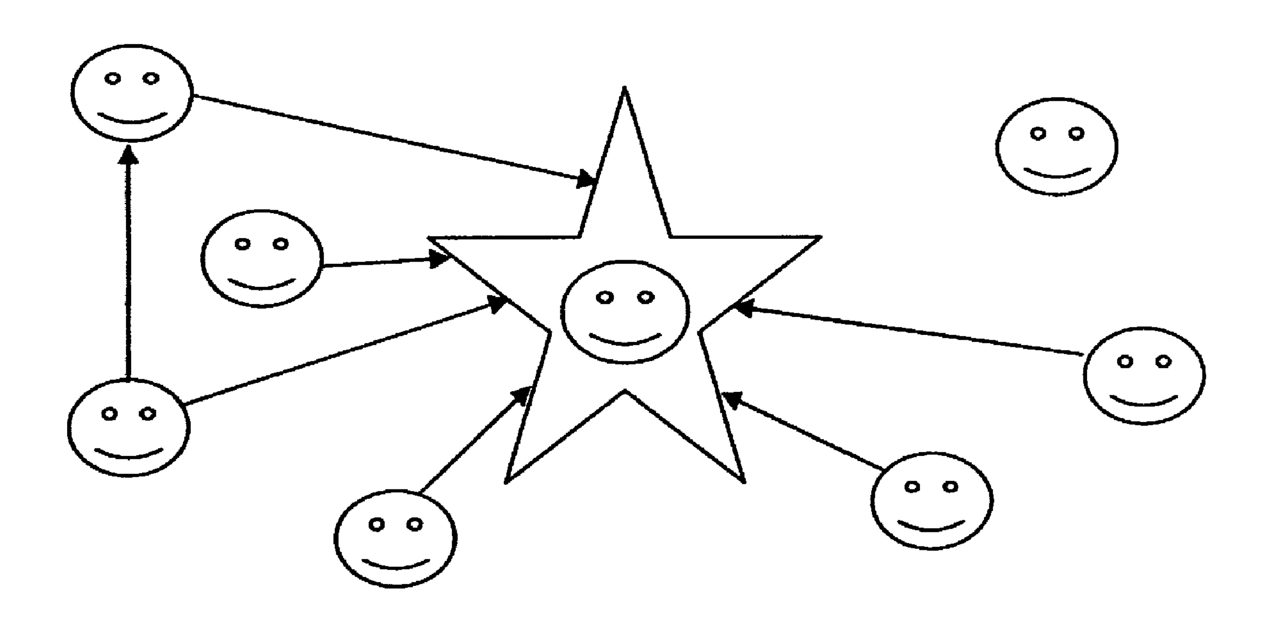
وعلى الرغم من ذلك فقد أوضحت الدراسات ان صدق مقاييس تمايز معاني المفاهيم وثبات درجاتها مرض بوجه عام. فالارتباط بين درجاتها ودرجات مقياس ثيرستون وليكرت تراوحت مابين 0,74 و0,82 في حين كانت معاملات ثبات هذا النوع من المقاييس باستخدام التجزئة النصفية مابين 0,83 و 0,91

كما أن هناك طرقاً أخرى عديدة توصلنا إلى نفس الهدف في قياس الاتجاهات مثلا الأسلوب الإسقاطي أو الاختبارات الاسقاطية من خلال استجابة الفرد لرسوم غامضة أو من خلال تصرفاته تجاهها وبالتالي معرفة اتجاهاته نحوها. فالفرد عندما يطلب منه القيام بدور يحبه تراه يتقنه ويقبل عليه برغبة زائدة والعكس صحيح أيضاً. كما ان محاولة سؤالنا الفرد عن الأشياء التي يحبها أو يكرهها يعطينا صورة واضحة عن اتجاهاته عن هذه الأشياء.

أسلوب قياس العلاقات الاجتماعية Sociometric Techniques

يهدف هذا الأسلوب دراسة التفاعل الاجتماعي بين أفراد جماعة معينة ومعرفة أنماط هذا التفاعل الذي يمكن أن يكون سلوكاً فعلياً أو مشاعر أو توقعات، حيث يمكن التوصل إلى معرفة ذلك عن طريق استخدام الاختبارات السوسيومترية Sociometric Questionnaires أو المقابلات الشخصية Interviews حيث يطلب من كل فرد تحديد الأفراد المفضلين لديه في القيام ببعض الأنشطة والأفراد الذين لا يريد مشاركتهم.

ويمكن استخدام التمثيل الشكلي للعلاقات الاجتماعية الذي اقترحه Moreno لتحليل نمط هذه العلاقات بين أفراد الجماعة كما هو موضح بالشكل الآتي:



تستخدم مثل هذه الأشكال عندما يكون العدد في الجماعة قليلاً أما إذا كان العدد كبيرا فإنه يستعاض عنها بما يسمى بمصفوفة العلاقات الاجتماعية وهي في صورتها البسيطة عبارة عن جدول يوضع في بعديه الرأسي والأفقي أسماء الجماعة، ونعبر بالرقم (1) عند وجود علاقة بين فردين أو عند اختيار أحد الأفراد للآخر، وبالرقم (0) إذا لم توجد العلاقة أو عدم الاختيار.

لقد اشترط مورينو عدداً من الشروط التي يجب توفرها في أداة القياس السيوسيومترية لكي تصبح صالحة للتطبيق والتحليل **وتتلخص في:**

1- سرية استجابات المفحوصين.
2- وضوح حدود جماعة الاختيار كأن تكون جماعة الفصل المدرسي مثلاً.
3- تحديد نوعية الموقف الاجتماعي بحيث لا يكون عاماً شاملاً ويحتمل أكثر من تأويل.
4- يجب أن يكون الموقف الاجتماعي حقيقياً وله صلة واضحة بالحياة اليومية لأعضاء الجماعة.
5- حرية الاختيار والرفض متروكة دون تحديد العدد حيث إن الفرد حر في اختيار أي عدد يشاء.
6- تنبيه الجماعة إلى أهمية اختياراتهم أو رفضهم لإعادة تنظيم الجماعة أو عند قيامها بنشاط معين.

ولبناء المقياس السوسيومتري ينبغي اتباع الخطوات الأساسية التالية:

1- اختيار الموقف الاجتماعي المناسب لكل جماعة مما يتصل بحياتها اليومية، مثل جماعة الفصل أو جماعة نشاط مدرسي معين.
2- صياغة السؤال السوسيومتري بطريقة صحية وسليمة من حيث مناسبة اللغة لمستوى العمر واستخدام الألفاظ ذات المفاهيم الواضحة وكذلك من حيث دلالته المباشرة على الموقف الاجتماعي دون تأويل.
3- إعداد تعليمات الاختبار السوسيومتري بحيث تكون سهلة وبسيطة ودقيقة يمكن فهمها دون تعقيد، وان تكون محايدة لا تحتوي على إيحاءات مباشرة أو غير مباشرة لاختيار فرد أو رفض آخر.
4- تطبيق الاختبار ومن ثم تحليله وذلك بحساب الدرجة السوسيومترية وبناء المصفوفة السوسيومترية واستخراج المعاملات السوسيومترية مثل معامل التأثير

ومعامل التفاعل النفس اجتماعي ومعامل ثبوت الجماعة ومعامل التماسك الداخلي للجماعة حيث يوجد صيغ رياضية لاستخراج جميع المعاملات السابقة.

معامل التأثير:

يستخدم عند الرغبة في اختيار الزعامات أو دمج الجماعات الصغيرة، ولمقارنة المكانة الاجتماعية لفردين أو أكثر.

$$\text{معامل التأثير} = \frac{\text{عدد الاختيارات الفعلية التي حصل عليها الفرد (ك)}}{\text{عدد أفراد الجماعة (ن)} - 1}$$

معامل التفاعل النفس اجتماعي:

يستخدم لمعرفة مراحل نمو الجماعة في فترات مختلفة، ولمقارنة المكانة الاجتماعية لفردين أو أكثر.

$$\text{معامل التفاعل النفس اجتماعي} = \frac{\text{المجمع الكلي للعلاقات الفعلية}}{\text{ن (ن} - 1)}$$

ن = عدد أفراد الجماعة

معامل ثبوت الجماعة:

يستخدم لمعرفة مدى تكامل وقوة الجماعة ومقاومتها للضغوط الهادفة لتعديل بنائها وتكوينها.

$$\text{معامل ثبوت الجماعة} = \frac{\text{عدد الأفراد المقاومين للتغير (ج)}}{\underset{\text{(قبل التغيير)}}{\text{عدد أفراد الجماعة (ن)}} + \underset{\text{(بعد التغيير)}}{\text{عدد أفراد الجماعة (ص)}}}$$

معامل التماسك الداخلي للجماعة:

ويهدف لمعرفة العلاقات السوسيومترية داخل الجماعة عندما تقع تحت تأثير جماعة خارجية، وقياس العلاقة بين هاتين الجماعتين.

$$\text{معامل التماسك الداخلي} = \frac{\text{م} (\text{ل} + \text{هـ})}{\text{ن} \times \text{د}}$$

م = عدد أفراد الجماعة الخارجية المؤثرين

ل = عدد العلاقات الداخلية الفعلية

هـ = عدد العلاقات الداخلية الفعلية

ن = عدد أفراد الجماعة الداخلية

د = عدد العلاقات التي تخرج من الجماعة الداخلية

وفيما يخص قضية الصدق والثبات لهذا النوع من المقاييس فيرى علماء القياس أنهما مازالتا محط نقاش وبحث.

تحليل المحتوى

يقصد بالمحتوى "هيكل من المعاني في صيغة رموز، وقد تكون الرموز لفظية أو صورية أو إشارية) وهذه المعاني هي التي تشكل لنا عملية الاتصال.

أما التحليل فيقصد به (تفتيت البيانات وتنظيمها في عناصر أساسية لغرض الحصول على إجابات للأسئلة التي آثارها البحث).

إن تحليل المحتوى هو طريقة لدراسة وتحليل وسائل الاتصال بأسلوب منظم وموضوعي وكمي لغرض قياس المتغيرات.

ويعود ظهور طريقة تحليل المحتوى إلى العقود الأولى من القرن التاسع عشر، فقد استخدمت هذه الطريقة لأول مرة في مجال تحليل المحتوى الصحف للتعرف على الأفكار والقيم والاتجاهات المختلفة، وكذلك استخدمت في ميدان الأدب والسياسة

واللغة. ومن ثم في مجال الدعاية والرأي العام. وقد استخدم هذا الأسلوب خلال الحرب العالمية الثانية استخداما واضحا من قبل بعض المؤسسات الحكومية في مجال الاتصال الجماهيري.

وبذلك يمكن تعريف تحليل المحتوى على إنه (وسيلة أو طريقة لتحليل وسائل الاتصال في مجالاتها المختلفة تحليلا علميا وبمنهجية محددة تتصف بالموضوعية تمهيداً لوصفها كماً وكيفا ").

خصائص تحليل المحتوى:

1- أنها طريقة استخدمت بنجاح ملحوظ في مجالات أخرى غير العلوم الاجتماعية كالصحافة والأدب والإذاعة وغيرها.

2- أنها تؤكد على مسألة معرفة وتقرير الآثار الناتجة عن الاتصال بين الأفراد، سواء أكان هذا الاتصال بشكل رموز أو معاني أو حديث إذاعي أو مقالة صحفية.

3- أنها طريقة تؤكد على المحتوى الظاهر للاتصال ولا هتم بالمقاصد الكامنة التي تحدد المحتوى.

4- أنها طريقة علمية تتسم بالموضوعية، أي بمعنى أن كل خطوة في عملية البحث ينبغي أن تنفذ على أساس قواعد وإجراءات موضحة بصورة لا قبل إلا الصراحة. أي معرفة التصنيف الذي سيلائم المادة المحللة، والمعيار أو المحك الذي يبغي استخدامه لكي نقرر وحدات تحليل المحتوى.

5- أنها تتسم بالتكميم، والتكميم لا يتطلب دائما إعطاء قيمة رقمية للأصناف بل يأخذ في بعض الأحيان صورة تقديرات كمية مثل (أكثر) (دائما)

6- شرط التنظيم والذي يؤكد على ضرورة استبعاد التحليلات الجزئية أو المتحيزة التي تحتوي على مجموعة عناصر اختيرت لدعم وجهة نظر المحلل فقط.

القيم في تحليل المحتوى:

إن أول بحث في القيم استخدمت فيه طريقة تحليل المحتوى تم في سنة 1947 من قبل (وايت) وكان يهدف من وراء ذلك إثبات إمكانية الوصول إلى درجة عالية من الموضوعية.

وتختلف عينة المحتوى في مجال القيم بحسب طبيعة الدراسات وأهدافها والشيء الذي نقوم بتحليله قد يكون عبارة عن (إجابات عن بعض أسئلة البحث، إصدارات من الصحف والمجلات، مواد كتب دراسية أو ثقافية أو اجتماعية أو سياسية، خطب وأقوال القادة السياسيين والاجتماعيين وغيرهم، برامج إذاعية أو تلفزيونية، نشاطات معينة عن طريق الملاحظة أو السلوك).

وتستخدم العشوائية في الكثير من الحوث لانتقاء نسبة من عينة المحتوى وخاصة في مجال التعرف على القيم السائدة في الصحف والمجلات. وعينة المحتوى يمكن أن تكون ممثلة إذا أخذت بنظر الاعتبار الأمور المهمة في تحديد عينة البحث والتي تتأتى من إدراك الباحث لطبيعة المحتوى الذي يقوم بتحليله والاتجاهات المؤثرة فيه، وعامل الزمن، والتغيرات التي يمكن أن تحدث في مجال القيم التي يحللها.

وحدات التحليل:

تعتمد وحدة التحليل على هدف أو أهداف البحث، طبيعة المادة التي يرغب الباحث في إخضاعها للتحليل، وتستخدم عادة في تحليل المحتوى وحدتان تحليليتان هما:

1- وحدة الترميز أو التسجيل (وهي اصغر جزء في المحتوى المحلل، يتم عن طريقها إحصاء ما يراد تشخيصه في ذلك المحتوى). إن وحدة الترميز أو التسجيل التي تستخدم عادة في بحوث القيم هي (الفكرة Theme) لأنها أكثر الوحدات ملائمة لأبحاث القيم والاتجاهات قياساً بالوحدات الأخرى كالكلمة والموضوع.

2- وحدة المضمون (السياق) فهي الهيكل المحيط بوحدة التسجيل، وقد تكون وحدة المضمون في البحث الفقرة والموضوع الذي تقع فيه الفكر.

ويمكن أن نتعرف من أبيات تحليل المحتوى على وجود خمس وحدات أساسية للتحليل وهي:

- الكلمة Word وهي اصغر وحدة تستخدم في تحليل المحتوى.
- الفكرة Theme
- الشخصية وتستخدم في تحليل القصص والتمثيليات وتاريخ الحياة.
- الموضوع Item ويعتبر أكثر استعمالا في تحليل المحتوى ويقصد به الوحدة الكاملة التي تقدم من خلال مادة الاتصال.
- مقاييس الزمن أو المسافة: إن بعض الدراسات تقوم بتصنيف المحتوى على أساس أقسام مادية مثل (حقل الانجات، أو الصفحة، أو السطر أو العبارة (بالنسبة للطباعة)، وقد تقاس بالدقيقة في الراديو.

وحدات التعداد:

استخدم التكرار في أكثر البحوث المتعلقة بالقيم كوحدة للتعداد للإجابة الكمية على هدف البحث واتفقت معظم الدراسات على إعطاء أوزان متساوية لكل من الوحدات في المحتوى المحلل.

خطوات التحليل:

تتم عملية التحليل على وفق الخطوات الآتية في مجال تحليل القيم:

1- يقرأ الموضوع أو القصة أو المقالة ككل، للتعرف على الفكرة الأساسية فيه (وحدة المضمون أو السياق)

2- قراءة وحدة العينة (الصفحة) وتحديد الفكر (وحدات التحليل) التي تتضمن قيماً وتمييز الفكر الضمنية عن الصريحة.

3- تعيين صنف القيمة في كل فكرة في ضوء أداة البحث (وحدة التسجيل).
4- تفريغ نتائج التحليل في جدول التحليل (استمارة التحليل) ويتم ذلك بإعطاء تكرار واحد عند ظهور أية قيمة من القيم.

قواعد التحليل:

إن إعطاء مجموعة واضحة وصريحة من التعليمات أو القواعد في التحليل يؤدي إلى تحديد العبارات بصورة أدق، كما يؤدي إلى ارتفاع نسبة الثبات. لذلك وضعت قواعد للتحليل وهي:

1- اعتبار العطف (theme) مستقلة في التحليل.
2- إن القيم سواء أكانت وسيليه أم غائية لاتهم بقدر ما يهم فيها أنها قيمة يبتغيها الفرد
3- اعتبار الأفعال المتكررة بصورة متتالية فعلا واحدا عند التحليل.
4- اعتبار التفسير الذي يقدمه المجيب لشرح ما يعنيه (ثيما) مستقلة.
5- استبعاد المقدمة من الإجابة.
6- إذا كانت القيمة سلبية فإنه يضاف حرف النفي (لا).
7- الاستمرار في قراءة صفحات أخرى من الموضوع، حينما لا تظهر أي دلالة قيمية في وحدة العينة إلى أن تكتمل الفكرة وتظهر دلالتها القيمية.
8- حينما تظهر فكرة لا تشير إلى أحد القيم المذكورة في تصنيف البحث، تصنف في مجال (القيم المتنوعة)
9- عندما تحتوي الفكرة الرئيسة والصريحة على فكر نوعية، فتعامل كل فكرة منها على أنها وحدة مستقلة للتحليل.
10- إهمال الصور والأشكال التوضيحية لعدم إمكانية خضوعها للتحليل حسب التصنيف.

الصدق في تحليل المحتوى

يعتمد الصدق في تحليل المحتوى على الأهداف المطلوب دراستها في البحث والتصانيف المستخدمة فيه. فعندما يستخدم الباحث تصنيفاً جاهزاً فقد يعتمد في تقديره لصدق المحتوى على أساس أن التصنيف الذي سيعتمده يفترض "الصدق" فيه وانه يمكن استخدامه في قياس وتشخيص القيم. ويبرر البعض هذا الافتراض في الحصول على الصدق بأنه تبين للباحث قدرة التصنيف على ذلك من خلال تحليل العينة الاستطلاعية أو تحليل عينة الثبات.

أما في حالة قيام الباحث ببناء تصنيف بنفسه لاستخدامه في مجال الإجابة على أهداف بحثه، فإنه يلجأ إلى اعتماد طريقة المحكمين التي من خلالها يتناقش الباحث مع الحكام وجهاً لوجه، أو من خلال عرض التصنيف على مجموعة من المختصين لغرض إقرار سلامة التعاريف والأفكار والأصناف الفرعية المدرجة تحت كل صنف من الأصناف الكبرى. وكذلك حذف وإضافة قيم جديدة تتناسب مع طبيعة البحث، ومن ثم تصنيفها إلى مجالات محددة بعد تسمية كل مجال منها، وحساب معامل الاتفاق على التصنيف و المجالات بعد ذلك.

الثبات:

لما كانت طريقة تحليل المحتوى تشترط الموضوعية، وأن تحقيقها يتطلب توفر صفة الثبات، والذي يعتبره (سكوت Scott) التعريف الإجرائي للموضوعية لذلك فإن الكثير من الباحثين لجأوا إلى استخراج ثبات التحليل. ويتأثر الثبات في تحليل المحتوى بعدة متغيرات أهمها:

1- خبرة المحلل ومهارته في التحليل.
2- نوع التصنيف ومدى وضوح أصنافه.
3- نوع وحدة التحليل.
4- مدى وضوح قواعد التحليل.
5- نوع البيانات المحللة.

إن كل ذلك يتطلب من الباحث أن يقوم بمجموعة من الإجراءات اللازمة للحد من الأخطاء التي تعمل على خفض الثبات. ويتم ذلك من خلال اعتماد الباحث على إجراءات ستامبل التي تؤكد على:

1- توضيح أصناف التصنيف وتعريفها بشكل جيد.
2- حسن اختيار وتدريب المحللين.
3- السير بالتحليل بطريقة واضحة.

وتتطلب هذه الإجراءات ما يلي:

1- إن يحاول الباحث زيادة مهارته وخبرته في التحليل عن طريق الممارسة العملية للتحليل لمجموعة من العينات وعرضها على مجموعة من الخبراء لإقرار سلامتها.
2- اختيار الأفراد الذين لهم الخبرة في هذا المجال كمحللين للثبات وتدريبهم على تحليل عينته وبشكل مستقل أحدهما عن الآخر، وعن الباحث في نفس الوقت، لغرض التأكد من سلامة التحليل، ومناقشة النتائج فيما بعد وتسوية بعض الخلافات التي قد تظهر عن طريق تعديل بعض قواعد التحليل.
3- الاستعانة بمن لهم خبرة في هذا المجال عن طريق تشكيل لجان تقوم بالتحليل بصورة جماعية وبحضور الباحث الذي يستمع إلى مناقشاتهم أثناء التحليل، ويسجل نقاط الاختلاف التي تثار خلال عمليات التحليل، فيستفيد بذلك في تحسين بعض التعاريف وتوضيحها وصياغة بعض قواعد التحليل.

أن ثبات تحليل المحتوى يستند على نوعين من الاتساق (Consistency) هما:

1- الاتساق بين المحللين: ويعني توصل محللين مختلفين إلى نفس النتائج عند استخدام نفس التصنيف والمحتوى وإتباعهما إجراءات التحليل نفسها.
2- الاتساق خلال الزمن: ويعني توصل محلل منفرد (أو مجموعة من المحللين) إلى نفس النتائج عند استخدام نفس التصنيف على نفس المحتوى في فترات زمنية مختلفة.

ولتحقيق الاتساق الأول، يختار الباحث أثنين من المحللين الخارجيين ممن لهم خبرة ودراية في هذا المجال بعد تدريبهما على عملية التحليل (لتحليل عينة من الحالات تسحب عشوائياً وبصورة منفردة).

أما الاتساق الثاني فيتم بتحليل عينة من الحالات المطلوب دراستها من قبل الباحث، وتحليلها مرة ثانية بعد مرور فترة (30) يوماً وهي فترة ذكرت في الكثير من دراسات القيم. ومن ثم يتم استخراج معامل الاتفاق بنوعيه عن طريق (معامل اتفاق سكوت) ويمثل إحصائياً بالآتي:

$$\text{معامل الاتفاق} = \frac{\text{مجموع الاتفاق الكلي بين الملاحظين} - \text{الاتفاق الناجم عن الصدفة}}{\text{أكبر اتفاق ممكن} - \text{الاتفاق الناجم عن الصدفة}}$$

ويذكر لنفلي أن معامل الثبات الذي يتراوح مابين (0,50–0,60) يعتبر مرضياً.

ويرى بيرلسون Berlson أن الثبات يرتفع عندما يستخدم (تصنيف بسيط) غير معقد ووحدة تحليل بسيطة. وعلى العكس من ذلك فإن الثبات يهبط عندما يستخدم تصنيف معقد ووحدة تحليل معقدة. وهذه بحد ذاتها مشكلة في طريقة تحليل المحتوى، لأن استخدام تصنيف بسيط يعطينا ثباتاً عالياً، غير أن معلوماته لا تكون غنية. وتعتبر قائمة (وايت) من التصانيف المعقدة، ولذلك لا يكون الثبات فيها عالياً.

التصانيف المستخدمة في التحليل:

تعتمد التصنيفات المستخدمة في تحليل المحتوى على هدف البحث وهناك اتجاه يشير إلى اختيار تصانيف جاهزة كتلك التي يطلق عليها Holsti بالتصانيف المعيارية، بعد التأكد من إيفائها بمتطلبات الدراسة وملاءمتها. بينما يشير الاتجاه الثاني إلى بناء تصنيف لغرض الدراسة ويرفض التصنيفات الجاهزة أما لعدم إيفائها بمتطلبات الدراسة، أو لتباين أهدافها، أو كونها غير مؤهلة لطبيعة المحتوى المحلل.

وهناك عددا من الدراسات قامت ببناء تصنيف خاص بها سواء بطريقة التحليل

القبلي أو البعدي. وهناك دراسات استخدمت تصانيف جاهزة ولكنها طورت عن التصنيف الأساسي لجعلها أكثر قدرة وملائمة للبيئة الجديدة.

1- بناء التصانيف ويتم هذا البناء بطريقتين هما:

أ- الطريقة القبلية: وتتلخص باستخدام نظام نظري مسبق يوضع كتصنيف للدراسة المقترحة يغطي ما يهدف إليه البحث ومحاولا تجميع كل القيم الممكنة في هذا المجال. وتتميز هذه الطريقة بارتكازها على أطار نظري يسهل عملية التفسير فيما بعد، كما تتميز بالثبات وسهولة التحليل الإحصائي.

ب- الطريقة البعدية: وتستمد التصنيفات في هذه الحالة من المادة المراد تحليلها بعد جمعها وحصرها بطرق متعددة. ويعتمد نجاحها على قدرة المحلل على أفراد الأفكار الهامة في المواد المحللة، وأن يعمد إلى إيجاد مخطط عملي لتصنيف البيانات.

2- اختيار التصنيفات الجاهزة:

إن الكثير من الدراسات المتعلقة بالقيم قد أخذت بتصانيف جاهزة للتعرف على القيم في مجالات متعددة، واستخدمتها وتعرفت على القيم المطلوبة. إلا أن هناك دراسات أخرى أخذت تصانيف جاهزة ولكنها طورت لجعلها أكثر وضوحا وملائمة للبيئة.

إن من التصانيف المشهورة في مجال القيم (تصنيف وايت) كما هو في صورته الأصلية وكذلك التصنيف المطور منه والذي أخذت به الدراسات العربية.

تصنيف وايت للقيم

يشير وايت إلى إنه قد اشتق تصنيفه مما كتبه علماء النفس الآخرون مثل (موري) و(توماس) و (سبرنجر) و (مكدوكل) و (ودورث) و (تولمان) في وصفهم "للحاجات والقيم والدوافع الإنسانية" فضلا عن قيامه بفحص القواميس السيكولوجية والسير

الذاتية والمقابلات الشخصية وقصص الأطفال والإعلانات ومحتوى المواد السياسية.

وصف القائمة:

تنقسم القائمة إلى مجالين رئيسيين هما:

1- مجال الأهداف Goals.

2- مجال معايير الحكم Standard Judgment.

وينضوي تحت كلا هذين المجالين عدد من "المجموعات القيمية" وتحتوي كل مجموعة قيمية بدورها على عدد من القيم "المنفردة" وفيما يلي مجال "الأهداف" والمجموعات القيمية والقيم المنفردة التي يتألف منها:

1- المجموعة الجسمانية: وتتضمن الطعام، الجنس، الراحة، النشاط، الصحة، الأمن.

2- المجموعة الاجتماعية: جب الجنس، حب الأسرة، الصداقة.

3- المجموعة الذاتية: الاستقلال، التحصيل، التقدير، احترام الذات، السيطرة أو التسلط، العدوان.

4- المجموعة الترويحية: الخبرات الجديدة، الاستثارة، الجمال، المرح.

5- مجموعة الأمن الانفعالي.

6- المجموعة العملية: القيمة العملية، الاقتصاد، الملكية، العمل.

7- المجموعة المعرفية: المعرفة (أو المعلومات).

8- المجموعة المتنوعة: السعادة

أما مجال "معايير الحكم" فإنه يتألف من المجموعات القيمية والقيم المنفردة التالية:

1- المجموعة الأخلاقية: الخلق، الصدق، العدالة، الطاعة، النقاء أو (الطهر)، الدين.

2- المجموعة الاجتماعية: الشخصية اللطيفة، التماثل أو التطابق، قواعد السلوك، التواضع، الكرم، التسامح، الاندماج في الجماعة

3- المجموعة الذاتية: القوة، التصميم، الذكاء، المظهر.

4- المجموعة المتنوعة: الحرص أو الانتباه، النظافة، الثقافة، التوافق.

وقد عرف وايت كل قيمة منفردة تعريفا "إجرائيا"، وضرب أمثلة في كيفية استنباطها من المحتوى وأعطى لكل قيمة رمزا يمثلها لتسهيل عملية التحليل.

نماذج من مقاييس القيم والاتجاهات:

مقياس الاتجاه نحو الناس العاجزين Attitude toward disabled people

وضع المقياس كل من Yuker & Block & Campbell عام 1960 ويهدف إلى قياس الاتجاه نحو الناس العاجزين بشكل عام يتكون المقياس من صورتين متكافئتين (A&B) كل صورة احتوت على (30) فقرة، نصف الفقرات تقريبا تتعلق بالتشابه والاختلاف في الصفات الشخصية بينما يعالج النصف الآخر أسئلة حول المعاملة الخاصة للعاجزين. تم تطبيق المقياس على عينة من طلبة كلية هوفسترا Hofstra لغرض استخراج الثبات. أما عينة الصدق فتم اختيار (248) عاجزا وأعطيت ستة بدائل للإجابة وهي (موافق جدا، موافق إلى حد ما، موافق قليلا، غير موافق قليلا، غير موافق إلى حد ما، غير موافق إطلاقا) واستخدمت لإيجاد الثبات طريقة التجزئة النصفية وتراوحت من 0,78–0,83 ومعامل التكافؤ للصورتين A والصورة B وتراوح مابين 0,41–0,83. وأوجد الصدق من خلال إيجاد الارتباط بين درجات المقياس مع درجات مقاييس أخرى مثل مقياس الرضا عن العمل وبلغ معامل الارتباط 0,46 وقائمة ادوارد للتفضيل الشخصي وبلغ معامل الارتباط 0,252.

واليك نماذج من فقرات المقياس

(الصورة A)

1- العاجزون غالبا ما يكونوا غير محبوبين
2- العاجزون أكثر انفعالا من الناس الآخرين

3- معظم العاجزين يشعرون بأنهم مثل الناس الآخرين
4- الأطفال العاجزون ينبغي وضعهم مع الأطفال الاعتياديين
5- معظم العاجزين أكثر وعيا بذاتهم من الآخرين
6- يرى العاجزون أنفسهم اجتماعيون

أما الصورة B

1- العاجزون هم دائما محبويين
2- معظم العاجزين يتزوجون وينجبون أطفالا
3- العاجزون غالبا ما يكونوا اقل عدوانا من الناس الاعتياديين
4- معظم العاجزين يتوقعون معاملة خاصة
5- يفضل معظم العاجزين العمل مع الناس العاجزين
6- لا يحتاج معظم العاجزين إلى رعاية خاصة

مقياس الاتجاهات نحو المتخلفين عقليا

اعد المقياس من قبل د. يوسف القريوتي من اجل استخدامه في مجال التعرف على تلك الاتجاهات سواء في مجال التطبيقات العملية أو الدراسات النظرية ذات العلاقة، تم بناء صورة أولية من المقياس اشتملت على (38) فقرة وفقا لطريقة ليكرت في قياس الاتجاهات، وجربت الصورة الأولية على (147) مفحوصا وبحساب مصفوفة الارتباطات بين الدرجة على كل فقرة وغيرها من الفقرات والدرجة الكلية تم اختيار الفقرات التي تحققت لها معاملات ارتباط عالية وعليه فقد اشتملت الصورة النهائية من المقياس على (22) فقرة تم استخراج ثلاثة دلالات من الصدق وهي المحتوى وصدق البناء والصدق التمييزي وتبين أن المقياس يتمتع بدرجة مقبولة من الصدق، واستخرجت دلالات الثبات للمقياس بطريقتين هي الإعادة للمقياس وبلغ معامل الثبات 0، 90 وحسب معامل آلفا وبلغ 0، 822 واتضح أن الاختبار يتمتع بدرجة عالية من الاستقرار والاتساق الداخلي ومن الأمثلة على

فقرات المقياس:

1- المتخلفون عقليا خطرون على السلامة العامة للمجتمع.

2- إن الفرد المتخلف عقليا في العادة يكون السبب في تعاسة أسرته.

3- لا يمكنني تناول الطعام مع شخص متخلف عقليا على مائدة واحدة.

4- بأي حال من الأحوال لا يمكنني العمل مع المتخلفين عقليا.

5- المتخلف عقليا إنسان مجنون.

مقياس القيم لجوردن البورت

من إعداد جوردن البورت وفيليب وجاردنر لندزي وعربه الدكتور عطية محمود هنا 1986 ويقيس الاختبار القيم الهامة المؤثرة في السلوك الإنساني. وحددت القيم بالنظرية والاقتصادية والاجتماعية والسياسية والجمالية والدينية، والاختبار في صورته الجديدة يتيح للفرد ان يطبقه على نفسه وان يستخرج النتائج وان يحدد قيمه الأساسية وهو في التوجيه التربوي والمهني والإرشاد النفسي.

ويتضمن القسم الأول بعض الأحكام والمشكلات التي يختلف الناس بشأنها فالمطلوب أن يبدي الفرد ما يفضله شخصيا بكتابة الرقم الصحيح في المربعات الموجودة إلى يسار كل سؤال ويتكون هذا القسم من (30) فقرة ومن الأمثلة على ذلك:

الفقرة (1)

إن الهدف من الجمعيات الدينية في الوقت الحاضر يجب أن يكون:

أ- تنمية الإيثار وعمل الخير

ب- تشجيع العبادة والصيام

الفقرة (2)

هل تفضل سماع سلسلة من المحاضرات عن:

أ- مقارنة أنواع الحكومات

ب- مقارنة الأديان ونشاطها

أما القسم الثاني فيتكون من (15) فقرة ولكل فقرة أربعة إجابات والمطلوب ترتيب الإجابات تبعا لتفضيلك الشخصي لكل منها وذلك بأن تكتب درجة لكل إجابة في المربع المناسب الموجود إلى اليسار وهذه الدرجات هي 1، 2، 3، 4 ومن الأمثلة على فقرات هذا القسم:

الفقرة1: إذا كنت في مجتمع يضم أصدقاء لك من نفس الجنس فما هو الموضوع الذي تفضل التحدث فيه؟

1- معنى الحياة

2- التطورات العلمية

3- الأدب

4- الاشتراكية

الفقرة 2: من الذي تفضله من هؤلاء الأشخاص:

1- فلورنس نايتنجيل (ممرضة مشهورة)

2- نابليون (قائد وسياسي)

3- هنري فورد (من رجال الأعمال والصناعة)

اختبار انتقاء القيم

وهذا الاختبار هو الأكثر استعمالا في ميدان الطب النفسي وذلك نظرا لبساطته وسهولة تطبيقه وكذلك سهولة استخلاص نتائجه بعد قيام الفاحص بإجراء فحصه العيادي، ويمثل الاختبار مجموعة من الأوراق (35) ورقة مثل أوراق العنب. كتبت على كل ورقة منها كلمة ذات مغزى فكري وجداني مثل: طعام، جبال، أولاد،

حياة، غابات، دراهم، سيارة، عائلة، ثروة، أزهار، أغنية، موت، متحف، هدوء، طبيعة، طفولة، موسيقى، بحر، صداقة، مسافرين، مسرح، نبيذ، حلويات، سينما، ربيع، فقير، .. الخ.

يطلب من المفحوص أن يقرأ أولاً كل الكلمات المكتوبة على الأوراق وبعدها يجب على المفحوص أن ينتقي خمسة أوراق (من ال35 ورقة) ويقدمها للفاحص ويعيد الكرة مرتين أخريين بحيث يعطي الفاحص (15) ورقة. في غضون ذلك على الفاحص أن يراقب بحذر جميع حركات وكلمات المفحوص، ومن اجل استخلاص النتائج من خلال هذا الاختبار الإسقاطي كما نلاحظ على الفاحص أن يعرف بأن المفحوص المنفتح اجتماعيا ينتقي الكلمات الملائمة لهذا الوضع مثل: صداقة، مسرح، ربيع، ... الخ في حين أن المنغلق اجتماعيا ينتقي كلمات مثل: هدوء، طبيعة، جبال، غابات، ... الخ على إنه في بعض الحالات المتقدمة للانغلاق الاجتماعي مثل الميل للوحدة المرافق لحالات الانهيار فإن المفحوص قد يقدم للفاحص خمسة أوراق حتى دون أن ينظر إلى ما كتب عليها من كلمات وهذا يعكس موقف المنهار الرافض للفحص وللعلاج النفسيين. وهذا الموقف ذاته يعكس الشخصية الشبه انفصامية.

استفتاء القيم

بني من قبل د. حامد عبد السلام زهران ود. جلال محمد سري من جامعتي عين شمس والأزهر عام 1985 يهدف الاستفتاء إلى دراسة القيم السلوكية الموجودة فعلا عندك في الوقت الحاضر، ويتسم ترتيب الاستجابة على كل عبارة حسب أولوية وجودها عند المستجيب كسلوك فعلي، وتبلغ عدد فقرات الاستفتاء (48) فقرة ومن الأمثلة على بعض فقرات الاستفتاء ما يلي:

الفقرة 1:

1- العمل على حل المشكلات الاجتماعية.

2- العمل على الحصول على الثروة واستثمارها في التنمية والإنتاج.

3- العمل على تشجيع الفن والابتكار الفني.

الفقرة 2

1- قراءة الصحف والمجلات مهتما بصفحة الاجتماعيات..
2- قراءة الصحف والمجلات مهتما بصفحة الاقتصاد والأسواق.
3- قراءة الصحف والمجلات مهتما بصفحة الفن والمعارض الفنية.

الفقرة 3

1- اختيار العمل في مجال الدين.
2- اختيار العمل في وظيفة ذات مسؤوليات كبيرة.
3- اختار العمل في البحث العلمي لاكتشاف الحقائق.

الفصل الحادي عشر

قياس الشخصية

يحتوي الفصل على:

- مقدمة
- وسائل التقرير الذاتي
- المقابلة
- تاريخ الحالة
- مقاييس التقدير
- تقدير السمات عن طريق الاستبيانات والقوائم
- الاختبارات الإسقاطية
- تصنيفات أخرى لقياس الشخصية
- المشكلات التي تواجه اختبارات الشخصية وقياسها

مقدمة

إن نظرة تاريخية إلى تطور القياس التربوي والنفسي نجد أن تلك المقاييس في مجال الشخصية قد ظهرت متأخرة إذا ما قورنت بغيرها من المقاييس في المجالات الأخرى في علم النفس. ونعتقد إنه يعود لأسباب عديدة منها إلى تعقد مجال الشخصية

ككل، وكثرة الأبعاد والمتغيرات التي تقضي دراستها في ميدان الشخصية.. إلى جانب ضعف نسبة الاتفاق بين الباحثين في هذا المضمار في تعريف الشخصية وتحديد أبعادها ومدى تأثيرها ومنهجية دراستها.. ووفقاً لما تقدم وجدنا أن أساليب قياس الشخصية قد تباينت وفقاً للنظريات التي تكمن وراء الطريقة المستخدمة في القياس،أو حسب مناطق الشخصية المراد دراستها كالسمات والأفكار، وحسب نمط المثير الذي يعرض على المفحوص كالورقة والقلم أو الصور أو حسب نمط الاستجابة المطلوبة كأن تكون مقيدة أو حرة،أو حسب ظروف الإجراء والتعليمات وطريقة التفسير والأهداف التي تخدمها الاختبارات وغير ذلك من التصانيف.

ويستند قياس الشخصية إلى عدة افتراضات أهمها:

1- إن الأفراد يختلفون فيما بينهم، وان دراسة الشخصية تتطلب إمكانية تقدير وقياس هذه الفروق.

2- إن معظم السمات السلوكية ثابتة للقياس وتتدرج في استمرارية تخضع له.

3- الشخصية تتميز بشيء من الثبات ولكنه ليس ثباتاً مطلقاً.

وليس غريباً أن يرى البعض أن القياس النفسي والتربوي هو قياس الشخصية بكل جوانبها وهم في هذا يستندون إلى فهمهم أو نظرتهم للشخصية باعتبارها تتضمن كل الجوانب العقلية أي المعرفية وغير العقلية التي تتضمن الجوانب المزاجية والاجتماعية أيضاً.

ووفقاً لما تقدم فإن المقاييس النفسية تقسم إلى قسمين رئيسيين هما:

أ- مقاييس الأداء المتميز بمعرفة ما يقوم به الشخص عادة.

ب- مقاييس أقصى الأداء بمعرفة أقصى ما يمكن أن يقوم به الشخص في ظروف استثارة ودافعية غير عادية.

وهناك ميل واضح لاستخدام لفظة مقياس أو استفتاء أو استخبار أو قائمة أو

سجل في قياس الشخصية بدلاً من لفظ اختبار، حيث توحي التسمية الأخيرة بوجود إجابة صحيحة وأخرى خاطئة كما يحدث في الاختبارات المعرفية أو العقلية كاختبارات الذكاء والقدرات، بينما يطلب من المختبر في قياس الشخصية أو مكوناتها أن يجيب عن الأسئلة أو العبارات أو القضايا أو المواقف الموجودة أمامه بما يشعر به وبما يفضله وبما يكرهه وأساليب سلوكه واتجاهاته نحو مواضيع خاصة. وعليه لا توجد إجابات صحيحة وأخرى خاطئة بل كل لها قيمتها ودلالاتها.

يقسم كرونباك أنواع مقاييس الشخصية بما يلي:

1- قياس الشخصية عن طريق التقرير الذاتي self-Report.

2- قياس الشخصية عن طريق أحكام الآخرين والملاحظات المنظمة Judgments and systematic observations.

3- قياس الشخصية عن طريق اختبارات الأداء performance tests.

يسعى مصممو مقاييس الشخصية وضع عبارات عن الأشخاص لترينا تجريبياً أنها مفيدة ودقيقة، ونحصل بموجبها على أنواع مختلفة من البيانات بعضها تكون غير شخصية وحتى بعضها لا تتطلب الملاحظة المباشرة للشخص الخاضع للدراسة، ففي قياس مركز العلاقات الاجتماعية لشخص ما قد لا نحتاج إلى الاتصال مع الشخص ولا إلى التاريخ الطبي والأسري على الرغم من أهمية وقيمة المعلومات التي نحصل عليها في دراسة حياة الإنسان، إنهم يميلون للاهتمام بتلك المعلومات التي تستند على الملاحظة المباشرة لسلوك الأفراد.. إن بحوثاً عديدة في مقاييس الشخصية تخضع للبرنامج التالي: الباحث لديه فرضية عن واحد أو أكثر من خصائص الشخصية يعتقد أنها ترتبط مع بعض جوانب السلوك فمثلاً يعتقد أن القلق والكبت يتحددان بطرق خاصة ويرتبطان مع النجاح في الزواج. لذا فهو يلاحظ جوانب مختارة من سلوك الفرد عن طريق المقاييس والمقابلات ومن نتائجها يستنتج المؤشرات لتلك الاستعدادات.. إن اعتماد الدقة في اتباع هذه الخطوات

يبدو تقود الباحث إلى مشكلات وتعقيدات حادة كصفة الاستقرار في السلوك والثبات في الدرجات لمؤشرات خصائص الشخصية، وفيما يتعلق بالاستقرار هل الفرد لديه نفس المؤشرات القائمة عند ملاحظة سلوكه في حالتين منفصلتين؟ ان الارتباط بين السمات الملاحظة وفي أوقات مختلفة هي التي تعطينا الاستقرار للخصائص المقاسة.

وهناك عامل آخر مهم في قياس الشخصية وهو الثبات للإجراءات المستخدمة بتقدير ثبات الشخص على مؤشرات الخصائص الشخصية...إن واضعي مقاييس الشخصية يكافحون من اجل استخدام مؤشرات خصائص الشخصية التي تتسم بالملائمة والموضوعية والثبات.

وعلى الرغم من الأهمية الفائقة إلى الثبات والاستقرار إلا أن القضية الجوهرية في القياس هي الصدق، ولكي يكون الاختبار صادق ينبغي أن يحرز على ثبات عال معقول، إلا أن الثبات العالي لوحده لا يظهر الصدق بوضوح.

إن الأنواع المتعددة للصدق تقود الباحث لاختيار الصدق المناسب للقياس، إذ أشار كرونباخ وميل cronbach &meehl أن الصدق التلازمي والصدق التنبؤي قد لا يكونا مناسبين في بعض بحوث الشخصية، فالباحث في الشخصية عليه أن يشغل نفسه بمشكلات يمكن حلها ببساطة عن طريق ارتباط درجات الاختبار الخاص مع إحدى مؤشرات الصدق وبصورة أدق هو يعالج المفاهيم والتضمينات التي تخترق مدى واسعاً من السلوك.

يقول ميلر: "إذا قورنت اختبارات الشخصية باختبارات الذكاء فإننا نجدها أقل ثباتاً لان درجة الشخص في اختبارات الشخصية قد تتغير تغيراً كبيراً من وقت لآخر، واقل صدقاً لان الصفات الشخصية نفسها قد تتغير من حيث معناها ومن كونها يجوب فيها أم لا ...... ".

وتعقيباً على ذلك فإن الباحثين يلجأون إلى وسائل عديدة لتحديد صدق

مقاييس الشخصية كطريقة معاملات الارتباط بالاعتماد على نتائج الاختبار بنتائج المحك المستخدم وطريقة الفئات المختلفة بعضها ببعض كأن نقارن نتائج الفصامين بالأسوياء في اختباراتهم، وطريقة المزاوجة بين نتائج الاختبار والصورة الإكلينيكية ...وتعتمد مقاييس الشخصية في صدقها على الصدق المنطقي والصدق التجريبي في معظم الأحيان وان كان بعضها قد تحدد صدقة عن طريق الصدق العاملي، وقد لوحظ إن معظم معاملات الصدق لاختبارات الشخصية متخصصة إذا ما قورنت بمعاملات الصدق لاختبارات القدرات...ويعتبر صدق الاختبارات الإسقاطية من أشق ما يقوم به الباحث فاختبار روشاخ مثلاً يحاول أن يمدنا بصورة كاملة عن الشخص وبذلك يصبح من الصعب قياس صحة الاختبار عن طريق اختبارات أخرى أو سلوك الفرد في مواقف محددة...وحاول البعض قياس صدقه على أساس تنبؤي ولكن لم يستطيعوا الحصول حتى الآن على نتائج ايجابية، كما أن اختبار تفهم الموضوع لم يكن أسعد حظاً من اختبار الروشاخ إذ لم يصل أحد الباحثين إلى نتائج قاطعة بشأنه.

أما فيما يتعلق بالثبات فهناك وسائل عديدة في مجال الشخصية أبرزها:

1- طريقة إعادة تطبيق المقياس على نفس المجموعة التي سبق أن طبق عليها، ففي قياس الشخصية نهتم بالتغير الذي يرجع إلى المقياس حتى نستطيع تحديد دقة المقياس المستخدم.

2- طريقة الصور المتكافئة للاختبار بتطبيق صورتين من صور الاختبار على نفس المجموعة واستخرج معامل الارتباط الموجود بين نتيجة الصورتين، وهذه الطريقة تعكس التغير الذي يعرف نتيجة انتقاء الأسئلة التي تتكون منها كل صورة من صور الاختبار.

3- طريقة تجزئة الاختبار (التجزئة النصفية) وهي تقوم على أساس تقسيم الاختبار إلى جزئين مع محاولة تساويهما في جميع النواحي.

طريقة تحليل التباين وتقوم هذه الطريقة على أساس أن أسئلة الاختبار مشتركة فيما بينها ومتجانسة

وسوف نقوم بتوضيح لمقاييس الشخصية وفقا للأسلوب كما يلي:

1- مقاييس التقرير الذاتي وتضم المقابلة الشخصية، تاريخ الحياة، اختبارات الورقة والقلم.

2- الاختبارات الإسقاطية كالصور وتداعي الكلمات وتكملة الجمل.

وسائل التقرير الذاتي Self Report

يؤكد هذا الاتجاه على العالم كما يدركه الفرد أو على العالم الذاتي للشخص بما في ذلك إدراك الفرد لذاته وان السلوك لا يفهم إلا في ضوء هذا العالم الذاتي الداخلي. ويذهب أصحاب هذا الاتجاه إن أغلب مشكلات التوافق التي يتعرض لها الفرد ترجع إلى إدراك الشخص للأحداث والأشخاص وليس لواقع هذه الأحداث والأشخاص في ذواتهم ... إن مفهوم التقرير الذاتي قيم ومتباين في المفهوم من عند كالتون يستعمل على إنه الوسيلة الوحيدة الممكنة الحصول على معلومات عن أمور وأحداث في عقل المفحوص، بيد إنه عند ستانلي هول يستعمل بهدف تجنب الجهد الشاق والوقت الطويل الذي يمكن أن يستهلك في الملاحظة المباشرة لسلوك المراهقين.

ويبدو أن هناك حدثان لهما دلالة خاصة في تاريخ اختبارات الورقة والقلم وبصور أدق في التقرير الذاتي احدهما كان من خلال تطوير اختبارات الذكاء في النصف الأول من هذا القرن وقد توصلوا إلى أن درجات اختبارات الذكاء لوحدها لا تحسب الفروق الفردية في السلوك ويبدو أنهم اقتنعوا بضرورة قياس الخصائص الشخصية أكثر من الذكاء،والحدث الثاني هو الحرب العالمية الأولى إذ أن التجنيد إلى العسكرية وبأعداد كبيرة من الجندين يستلزم بعض أنواع من الإجراءات المختارة لتحديد الرجال الذين شخصياتهم تمتاز بسوء التوافق.

لقد أوضح البورت أن أفضل ملاحظ لشخص ما هو الفرد ذاته، فإن الوصف الذاتي أو التقرير الذاتي يظهر لنا بيانات مناسبة عن شخصية الفرد على الرغم من اعتراض البعض عليها إزاء الصدق...هناك أدلة جيدة أن التقرير الذاتي ينبغي أن يأخذ بعين الاعتبار الدقة، والبحوث الحديثة في الشخصية تكرس جهودها للكشف عن مؤشرات النزعة الشخصية في وصف الذات.

وتوجد في حقل الشخصية أساليب عديدة لقياس التقرير الذاتي والتي طورت لتكون مفيدة بصورة خاصة في دراسات الشخصية وفي دراسات التقبل الاجتماعي من حيث تستخدم تقديرات الرفاق بصورة واسعة كما أن مقياس التقرير الذاتي منتشرة في دراسات علم النفس المرضي وأثار علاج الأمراض النفسية كمعرفة استجابة المرضى تجاه المواقف التي يجبونها في داخل المستشفى،واستجابتهم تجاه العلاج وأعضاء هيئة المستشفى.

ومن وسائل التقدير الذاتي هي ما يأتي:

1- المقابلة:

هي واحدة من أكثر أساليب القياس استخداماً من قبل السايكولوجيين والرؤساء وموظفي إدارة الأفراد والمعلمين، إن صدق المضمون في المقابلة يمكن الحصول عليه من خلال أحداث المقابلة والتي تكون أكثر وضوحاً عند الحديث باعتبارها مؤشرات موضوعية كانت ذات دلالة ارتبطت بحياة الأفراد إذ أن محتوى المقابلة يؤثر في اتجاهات وتوقعات ودوافع وإدراك وتعبيرات المقابلين (الأشخاص الذين يتعرضون إلى المقابلة) والاهم من ذلك أسلوب التفاعل المعتمد بين المقابل والمقابل.

إن المقابلة أسلوب لتقييم الشخصية شأنها في ذلك شأن الأساليب الأخرى، وهناك شيء من التشابه بين أسلوب سبر الغور عن طريق المقابلة والأسئلة التي توجه في اختبارات الشخصية أو مقياس الاستجابة للمفحوص في ضوء ما يقوله

الأخير كما يمكن أن يغير أسلوبه وأسئلته بما يلائم المفحوص .... وقد تكون المقابلة التي يعتمدها غير مقننة باعتبارها عرضية طارئة، أما المقابلة المقننة فإن القائم بالمقابلة لا يخرج عن الأسئلة التي وضعها مسبقاً وفي حالات أخرى يخرج عن بإعطاء الحرية لنفسه في وضع فقرات أخرى في المقابلة.

2- تاريخ الحالة case history

وتتلخص بجمع معلومات عن حياة الفرد منه ومن أهله وأقاربه والمتصلين به، والأساس في هذه الوسيلة افتراض شخصية الفرد الحالية ليست إلا مرحلة في عملية تطور مستمرة، وأنها نتائج ما مر به من خبرات وبالتالي فتاريخ الحياة يحدث بمؤشرات عن الخبرات التي يمر بها الفرد والتي ساهمت في تشكيله بقالبه الحالي.

3- مقياس التقدير Rating scales

يظهر الشخص بواسطة مقاييس التقدير استجابة مع أي فقرة باختيار رقم واحد من الاختيارات التي تبدو أنها أكثر ملائمة في وصف الفقرة، وهي في هذا مشابهة إلى فقرات الاختيار الإجباري عدا أن مقاييس التقدير تمثل درجات للخصائص المراد دراستها.

مثال:

- أشعر بالغضب تجاه الآخرين.
- دائما غالبا بعض الأوقات نادرا أبدا

وتستخدم هذه المقاييس في البحث والمواقف العملية بقياس رد الفعل عند الأفراد والتي تستخدم من قبل المشرفين والممرضين والأطباء ومعلمي التلاميذ... وتستخدم مقاييس التقدير عادة لقياس سمات كثيرة كالزعامة والأمانة والتعاون والمواظبة والكرم والغش....... الخ والقياس يتضمن عادة سمات تقدير بطريقة

ودية ويتوقف نوع السمات التي نقوم بتقديرها على الهدف الذي يوضع المقياس من اجله......وقد أكد فريمان على مبادئ عامة في بناء مقاييس التقدير منها:

1- تحديد السمة بوضوح.
2- تحديد درجة السمة إذ ينبغي تقديرها على مقياس متدرج من (5–7)درجات.
3- ثبات المقياس يعتمد على مدى تغير تقديرات الحكام.
4- تحديد صدق مقاييس الصدق إذ أن الوسائل العادية لتحديد الصدق قد يصعب استخدامها بالنسبة لمقاييس التقدير، وصدق مقاييس التقدير يفترض أن تقوم على فهم الحكام لمعاني السمات المراد تقديرها ومدى دقتهم في تقديرها.
5- السمات الظاهرية أكثر ثباتاً في التقدير في السمات الخفية أو الضمنية.
6- ذكر درجة الثقة في التقدير.

ومن ابرز الأنواع الشائعة في مقاييس التقدير ما يلي:

أ – مقياس التقدير الرقمية: يتحدد القائم بالتقدير قيمة عددية أو رقمية لكل سمة من السمات المراد تقديرها لدى الفرد. وفي المألوف عادة أن نجد المقياس الذي من هذا النوع يرتبط بصورة وصفية توضح للحكم على الأوزان الرقمية.

ب– مقياس التقدير البيانية: كتحديد الدرجات أو المستويات المعدة للسمة على نقطة معينة من خط مستقيم ويضع الحكم علامة على الموضع الذي اختاره للدلالة على السمة المراد تقديرها لدى الفرد بين الطرفين المتباعدين فالحكم هنا يضع علامة أو نقطة على الخط بدلاً من وضع درجة أو قيمة رقمية مثال:

السمة المراد قياسها: الاتجاه نحو الآخرين

- مشاكس وعنيد وغير متعاون ( )
- من الصعب هنا العمل معه ( )
- عادة لبق ومتعاون ويضبط نفسه ( )

- متعاون دائما ( )
- عامل أساسي في التعاون ورفع الروح المعنوية ( )

ج- مقياس الرتب: تستخدم بالنسبة للأشخاص الذين يوجدون داخل مجموعة واحدة ويراد معرفة وضعهم النسبي الواحد منهم للآخر.

د- مقياس تقدير قائمة المراجعة: حيث يكون المراد معرفة ما إذا كانت سمات معينة موجودة أو غير موجودة لدى الفرد. فمن الممكن استخدام ما يعرف باسم قائمة المراجعة وتتألف القائمة عادة من عدد من العبارات يعلم الحكم على الفقرة التي تنطبق على الفرد المراد.تقدير السمة عنده وأحياناً تعطى له تقديرات.

ومن مقاييس التقدير المعروفة: قوائم تقدير هاجرتي وويكمان وهي خصصت للكشف عن مشكلات السلوك والنزعات المشكلة ودراستها لدى الأفراد ابتداء من الحضانة حتى المرحلة الثانوية والقائمة (أ) عبارة عن تقدير لمشكلات السلوك وتحوي (15) نوعاً أو مصدراً لمشكلات السلوك كمشكلات الكلام والخروج على النظام وكل مشكلة منها تقدر من(1-4) درجات حسب تكرار حدوثها أما القائمة (ب) فهي مقاييس بيانية لـ (35) سمة مصنفة حسب أنواع أربعة عقلية وجسمية ووجدانية واجتماعية..... وتقدر هذه السمات وفق مقاييس من خمس فقط.

وهناك مقاييس أخرى منها مقياس فاينلاند للنضج الاجتماعي، ومقياس تقدير توافق التلميذ ولعل ابرز الانتقادات الموجهة إليها وجود التحيز لدى الحكام، وخطأ التفاعل بين الحكم والمقدر.

4- تقدير السمات عن طريق الاستبيانات والمقاييس:

إن البدايات الأولى لهذا النوع من المقاييس إلى ودورث حيث وضع قائمة للبيانات الشخصية Wood Worth personal data sheet وهو الأساس لعدد كبير من استفتاءات التوافق ولو أنها لا تتصف بالعمق في تلك المراحل السابقة، ولعل جهود كلفورد واضحة في هذا الميدان حيث اخضع نظريات الشخصية تماماً لبحث إحصائي عن أبعاد يمكن أن تلخص الشخصية، وقد أدت الارتباطات الداخلية بين عناصر الاختبارات إلى أن اقتراح كلفورد مثلاً أن الانطواء يمكن أن نقسمه إلى انطواء اجتماعي، انطواء فكري، واكتئاب........ أما الاتجاه الحديث في قياس الشخصية هو تحديد المفاهيم والتكوينات النظرية على أساس نظرية الشخصية تم إعداد العناصر في الاختبارات بصورة تؤدي إلى الحصول على معلومات عن هذه المفاهيم والتكوينات، وكان لنظرية يونخ عن الشخصية تأثير في الدراسات الأولى عن الانطواء، والتي تطورت باختبار مايرز - يرجز، كذلك مقياس التفضيل الشخصي لادواروز والذي اشتق من نظرية موراي عن الحاجات الخمس عشرة.

ويتكون مقياس أدورز من (210) زوجاً من العبارات وعلى المفحوص اختيار عبارة من كل زوج على أنها أكثر تمثيلاً لخصائص شخصيته مثال:

أ - احب أن أتحدث عن نفسي مع الآخرين.

ب- أحب أن أعمل لهدف معين حددته بنفسي.

ويمكن تصنيف هذه المقاييس إلى صنفين أساسيين هما:

1- المقاييس الأحادية البعد وتقيس سمة واحدة من أبعاد الشخصية كالانطواء والارتباط.

2- المقاييس المتعددة الأبعاد وتستخدم في قياس أكثر من سمة التي تكشف عن أهم سمات الشخصية التي تميز الأفراد بعضهم عن البعض.

ومن ابرز المقاييس الأحادية البعد:

أ- اختبار ودورث:

ويعد من أقدم الاختبارات العملية في قياس الشخصية ومنه استحدثت فقرات كثير من الاختبارات الأخرى التي وضعت بعد ذلك، وقد وضع وطور هذا الاختبار خلال الحرب العالمية الأولى حيث طلبت الهيئات المسؤولة في الجيش الأمريكي مساهمة علماء النفس في التعرف على الحالات غير الصالحة عقلياً من الجندين لاستبعادها عن ميدان القتال.... يتألف المقياس في صورته الأخيرة من (116) سؤالاً يجيب عليه المفحوص (بنعم) أو (لا) وقد مر وضع الاختبار بمراحل خمس هي:

1- وضع قائمة من الأسئلة مكونة من 200 سؤال اعتقد أنها تكشف عن المرض النفسي وقد استخدم من التعليقات الدارجة أو من ما هو موجود في الكتب.
2- طبقت هذه القائمة على مجموعة صغيرة من طلاب جامعة كولومبيا.
3- راجع إجابات الطلاب واستبعد منها الفقرات التي أجاب عنها أكثر من 25٪ من الطلاب إجابات عصبية.
4- تطبيق هذه الأسئلة على 1000 شخص عادي مختارين عشوائياً.
5- طبق نفس معيار الأبعاد السابقة الذكر على الأسئلة يتبين عدم صلاحياتها.

ويعتقد ودورث أن متوسط عدد الإجابات العصابية للعصابيين على هذا المقياس تقع بين (30–40) إجابة بينما يبلغ متوسط عدد هذه الإجابات بالنسبة للعاديين حوالي 10 إجابات من هنا نرى أن احتمال ارتفاع الدرجة (3) هذا المقياس تعني الكشف عن العصاب ومن أمثلة فقرات الاختبار:

1- هل تحس عادة بالصحة والقوة؟
2- هل تنام نوماً هادئاً عادة؟
3- هل تفزع من نومك أثناء الليل كثيراً؟
4- هل ينتابك الكابوس؟

ب- اختبار الشخصية لثرستون وثرستون:

وهو من وضع لويس ثرستون بحيث يتكون من (222) سؤالاً يجيب عنها المفحوص (بنعم) أو (لا) أو (لا أدري) ويهدف الاختبار كما يقول ثرستون إعطاء ميل ثابت نسبياً عن النزعات العصابية لدى طلاب الجامعات وقد مر المقياس بالخطوات الآتية:

1- جمع قائمة العبارات
2- طبع القائمة
3- تحديد الاستجابات على أساس قبلي
4- تطبيق الاختبار على مجموعة من الأشخاص
5- تحليل الفقرات
6- وضع المعايير

لقد جمع ثرستون أكثر من (600) فقرة من مصادر متعددة وطبعت كل فقرة على كل بطاقة مستقلة وصنفت إلى مجموعات عدة ثم أعيد ترتيبها وطبعت ثم اختصرت أخيراً ووصل عدد فقراتها 223 فقرة ...لقد حاولا التأكد من ملائمة هذه الأوزان التي وضعت بصورة قبلية تماماً بتحليل استجابات خمسين شخصاً الذين حصلوا على تقديرات عالية في هذا الاختبار، والخمسين شخصاً الذين حصلوا على اقل تقديرات فيه، ووجد أن الطلبة الأكثر عصابية كانوا يختارون الاستجابة الدالة على انحرافات عصابية في كل سؤال على حدة.

من فقرات المقياس:

- هل تجد صعوبة في أن تبدأ الحديث مع شخص غريب آخر لا تعرفه؟
- هل يجرح الناس شعورك عادة؟
- هل تسرح كثيراً في أحلام اليقظة؟

أما اختبارات الشخصية المتعدد الأبعاد منها:

1- قائمة بيرزويتر لقياس الشخصية: وهو من أوسع اختبارات الشخصية ويتكون من (125) سؤالاً تعرض لعدد كبير من الموضوعات التي تكشف عن سمات وخصائص الشخصية ومن ابرز مركبات الشخصية التي تم تناولها هي: مقياس الميول العصابية والذين يسجلون قياساً عالياً في هذا المقياس يميلون إلى عدم الثبات الانفعالي أما الذين يسجلون قياساً منخفضاً فهؤلاء على درجة حسنة من التوازن الانفعالي.

2- مقياس الاكتفاء الذاتي، فالذين يسجلون رقماً عالياً في هذا المقياس يفضلون أن يكونوا لوحدهم ويميلون إلى تجاهل نصائح الآخرين، أما الذين تكون درجاتهم منخفضة يكرهون الوحدة وكثيراً ما يسعون إلى طلب النصح.

3- مقياس الانطواء والانبساط والذين يحصلون فيه على درجات عالية يميلون إلى أن يكونوا انطوائيين خياليين يعيشون داخل نفسهم.

4- مقياس السيطرة والخضوع والذين يسجلون فيه قياساً عالياً يميلون إلى السيطرة على غيرهم في مواقف المواجهة أما الذين سجلوا قياساً منخفضاً فيميلون إلى الخضوع.

5- مقياس الثقة بالنفس والذين يحصلون على درجات عالية في هذا المقياس يميلون إلى الشعور بذاتهم بشكل منزعج ولديهم مشاعر الدونيه أما الذين يسجلون قياساً منخفضاً منهم فهم واثقون بأنفسهم تماماً.

6- مقياس الروح الاجتماعية فالذين يحصلون على درجات عالية على هذا المقياس يميلون إلى حب الاستقلال والوحدة والعزلة أما الذين يسجلون في قياس منخفض يبدون رغبة متمثلة إلى حب التجمع والاجتماع.

الاختبارات الإسقاطية

تعتمد الاختبارات الإسقاطية على المنهج الذي يحاول أن يصل إلى الكشف عن الشخصية عن طريق ما يسقطه الفرد على المثيرات الحسية من معان أو أشكال أو صور أو ما يؤكده فيها من نواح تتعلق بالون...وهو يقوم على افتراض أن ما يراه الفرد يرتبط بشخصيته كما يربطه بالمثير الذي يسببه.

تعتمد الاختبارات الإسقاطية أساساً على مفهوم الإسقاط الذي يتباين من منظر وباحث لأخر إذ يعرفه فرويد واليه الفضل الكبير في هذا المضمار هو احد العمليات الدفاعية التي يعزو بها الفرد دوافعه وإحساساته ومشاعره إلى الآخرين أو إلى العالم الخارجي، ويعد هذا بمثابة عملية دفاعية يتلخص بها الإنسان من الظواهر النفسية غير المرغوب فيها والتي أن بقيت سببت الألم للإنسان أما عند فرانك حينما وصف الوسائل غير المباشرة في دراسة الشخصية التي تهدف إلى الوصول بالفرد إلى أن يقدم تقييماً لصفاته دون أن ينتبه إلى أن يقوم بذلك.

وتمتاز الاختبارات الإسقاطية بما يلي:

1- الموقف المثير الذي يستجيب له الفرد غير متشكل وناقص التحديد والانتظام وبدوره يقلل من التحكم الشعوري للفرد في سلوكه بشكل يترتب عليه بسهولة الكشف عن شخصيته.

2- أن الفرد يستجيب للمادة غير المتشكلة التي تعرض عليه دون أن تكون لديه أية معرفة عن كيف أو من آية جهة سوف يتم تقدير هذه الاستجابات فدلالة المنهج غير معروفة لدى الفرد.

3- أنها تمثل نزعة من جانب الفرد ليعبر عن أفكاره ومشاعره وانفعالاته ورغباته في تشكيل المادة غير المتشكلة نسبياً.

4- أنها لا تقيس نواحي جزئية أو وحدات مستقلة تتألف منها الشخصية في مجموعها

بقدر ما تحاول أن ترسم صورة عن الشخصية ككل ودراسته مكوناتها وما بينها من علاقات ديناميكية.

وتقسم الطرق الإسقاطية إلى:

1- الطرق التكوينية أو التنظيمية : التي تتطلب من المختبر أن يفرض على المادة المعروضة وهي عادة غامضة أو قريبة من الغموض نوعاً من التكوين أو التنظيم كما في اختبار بقع الحبر لرورشاخ.

2- الطرق البنائية والإنشائية: والتي تتطلب من المختبر أن ينظم المواد المحددة الحجم كما في اختبار مجموعة اللعب، واختبار تكوين القصص المصورة.

3- الطرق التفسيرية: يتطلب من المفحوص تفسيراً لأشياء يجد فيها معنى شخصياً أو انفعالياً كما في اختبار تفهم الموضوعات قبل اختبار تفهم الموضوع.

4- الطرق التفريغية: إذ يتيح للمختبر أن يستعيد ويستخلص في انفعالاته كما في طريقة اللعب العلاجي بعمل الدمى وتحطيمها أو تشويهها وهي الطريقة التي ابتدعها "ليفي".

5- الطرق التحريفية: وهي الطرق التي تعطي صورة عن الشخصية عن طريق التحريف أو التغيير الذي يحدثه المختبر في أساليب الأفعال كأساليب الكلام المعينة.

الصدق والثبات في الاختبارات الإسقاطية

مشكلة الصدق والثبات في الاختبارات الإسقاطية من المشكلات الأساسية التي واجهت علم النفس الإكلينيكي،والبحوث العديدة لم تصد إلى حل قاطع وان كانت المحاولات تسير في سبيل إيجاد حل لها. وهناك خلاف ظاهر بين علماء النفس حول قياس صدق وثبات هذه الاختبارات ففريق من علماء النفس الإكلينيكي يؤمنون أن هذه الاختبارات تزود الباحث بمعطيات هامة عن ديناميات الشخصية

وان مسألة الثبات والصدق لا تعتبر مشكلة حقيقية وان الغرض الذي وضعت من اجله غرض إكلينيكي. كما يذهب نوتكات أن الاختبارات الإسقاطية لم توضع أساساً لقياس سمة واحدة كما أنها لا تقيم بطريقة آلية إذ أن الاختبار نفسه لا يتطلب الصدق بقدر ما يتطلب تأويله ذلك، فليس للاختبار الإسقاطي من معنى محدد قبل أن يتم تأويله. لقد استخدمت وسائل عديدة من اجل الحصول على الصدق والثبات منها.

1- قام رابابورت وجيل وشاتر بتحقيق صحة بعض الاختبارات التي تستخدم لتشخيص الأمراض العقلية عن طريق انتقاء مجموعات تجريبه تتكون من فئات إكلينيكية مختلفة من الذهانيين كحالات الاكتئاب ثم مجموعة ضابطة للمقارنة تتكون من (50) رجلاً من رجال المرور باعتبارهم أسوياء وبمقارنة استجابات المجموعات أمكن اختبار صدق العلامات المختلفة التي يستدل بها على المرض العقلي وهنا يعد الصدق صدقاً تلازمياً.

2- نسأل شخصاً يعرف المفحوص حق المعرفة فإذا كانت التأويلات التي حصلنا عليها من الاختبار تصدق فعلاً على هذا الشخص.

3- طريقة المضاهاة بإيجاد علاقة كمية بين نواحي نوعية في الشخصية كأن نضاهي بين تقريرين لباحثين مختلفين عن شخص واحد.

لقد تعرضت الاختبارات الإسقاطية إلى نقد كبير لها باعتبارها ذاتية وليست موضوعية وأن ثباتها وصدقها غير موثوق به وان قدرتها على التمييز بين الحالات السوية وغير السوية ضعيفة للغاية ومشكوك فيها.

نماذج من الاختبارات الإسقاطية:

1. الاختبارات التي تستخدم اللغة كمثير:

أ - اختبار تداعي الكلمات:

استخدمه جالتون كوسيلة لدراسة العمليات العقلية سنة 1879، 1883 كما

استخدمه بعض علماء النفس التجريبي مثل فونت وكاتل .... لقد وضعت قوائم كلمات كثيرة ومتعددة لدراسة التداعي أشهرها تلك التي وصفها يونج وتتكون من 100 كلمة اختيرت خصيصاً للكشف عن العقد ثم قائمة كينت وروزانوف وتتكون من 100 كلمة تجنبا فيها الإشارة إلى الكلمات المشبعة بالناحية الانفعالية والتي امتلأت بها قائمة يونج ومن القوائم التي تعد في نظر الكثيرين من أفضل القوائم تلك التي وضعها رابانورت وجيل وشافر وتتكون من 60 كلمة تتضمن مجالات متعددة كالأسرة والعدوان والدلالات الجنسية المتنوعة كما تمس مجالات فكرية.... وتقدم قائمة الكلمات في العادة شفوياً كما إنه يجرى فردياً لما يتطلبه الأمر من زمن الرجع لكل كلمة وملاحظة حركات المفحوص وإشاراته ...إن هذه القائمة تمدنا بمعلومات عن مجالات الاضطراب الانفعالي والتشخيص الإكلينيكي والكشف عن الجريمة.

ب- الاختبارات التي تستخدم الصور والأدوات كمثير:

من أبرزها اختبارات تفهم الموضوع تات TAT وتستخدم في أعمال العيادات النفسية وفي دراسة الشخصية وتدور فكرته تقديم عدد من الصور الغامضة نوعاً ما ودعوة المفحوص إلى تكوين قصة أو حكاية تصف ما يدور في الصورة وتتحدث عن أحوال الأشخاص والأحداث التي تجري فيها. ثم يقوم الفاحص بدراسة ما يقدمه المفحوص من قصة ويحاول أن يكتشف منها ما يحتمل في نفسه ميول ورغبات. إن أول من فسر هذا الاختبار هو هنري موراي وزميله مورجان سنة 1935 عن طريق فحصه الاخيلة والأوهام. إن الفكرة التي يقوم عليها هذا الاختبار هي أن القصص التي يعطيها المفحوص تكشف عن مكونات هامة في شخصيته على أساس نزعتين:

- الأولى نزعة الناس إلى تفسير المواقف الإنسانية الغامضة عما يتفق وخبراتهم الماضية ورغباتهم الحاضرة وآمالهم المستقبلية.

- الثانية نزعة كثير من كتاب القصص إلى أن يعترفوا بطريقة شعورية أو لاشعورية الكثير مما يكتبون من خبراتهم الشخصية ويعبرون عما يدور بأنفسهم من مشاعر ورغبات.

يتكون الاختبار من عشرة صور تقدم للمفحوص الواحدة بعد الأخرى يطلب منه تكوين حكاية أو قصة عن كل صورة منها وبعض الصور خاصة بالصبيان B وبعضها خاصة بالبنات G وبعضها خاص بالرجال M وبعضها خاص بالنساء F. وتعطى الصور وفق ترتيب محدد تشير إليه الأرقام المسجلة على ظهر البطاقة وتشير الحروف الأبجدية المكتوبة إلى جانب الرقم إلى نوع الشخص الذي تقدم إليه البطاقة ذكراً كان أم أنثى.

إن إجراء الاختبار ليس مشكلة بل المشكلة في تفسير المادة التي يعطيها المفحوص وقد أشار موراي أن تفسير الاختبار يتطلب الدقة والخبرة..... وقد صممت عدة استمارات لرصد وتحليل القصص أشهرها استمارة بلاك التي تتضمن ما يلي:

1- الموضوع الرئيس للقصة
2- البطل الرئيس فيها
3- الحاجات الأساسية للبطل
4- نظرة المفحوص البيئية
5- صور الشخصيات المختلفة في نظرة
6- أنواع الصراعات ذات الدلالة
7- طبيعة القلق عنده
8- الحيل الدفاعية الرئيسة لمواجهة أنواع الصراع والمخاوف
9- شدة الذات العليا
10- تكامل الذات

وهناك اختبار آخر لبقع الحبر وهو من وضع هرمان رور والذي توصل إلى وضع اختباره المشهور بعد سلسلة طويلة من التجارب التي أجريت على عديد من

يقع الحبر.... ويتكون هذا الاختبار من عشرة صور كل صورة منها متماثلة الشكل على نحو ما يحدث حين نلقي بقعة حبر كبيرة على ورقة بيضاء ثم تطبق الورقة ونضغط عليها فتخرج أشكالاً مختلفة متماثلة مع ذلك. وقد استخلص رورشاخ هذه الصور العشر من بين مئات الصور لأنها أكثر البطاقات قدرة على التمييز بين الحالات التي أجراها خمسة من هذه الصور تكون بدرجات مختلفة الظلال وصورتان أخريان من لونين اسود واحمر، أما الثلاثة الأخرى فتتكون من ألوان متعددة غير الأسود.

عند إجراء الاختبار لابد مراعاة جملة شروط منها عدم وجود شخص ثالث وان تكون الغرفة هادئة وجلسة الغموض مريح. يبدأ الاختبار عادة بتوضيح الطريقة التي عملت بها البطاقات...ثم تعطيه الفرصة للاستجابة ويبدأ الفاحص بتسجيل استجابات المفحوص على نحو ما يرويه لسانه وقد يستعين الفاحص ببعض الرموز والإشارات التي توضح موضع البطاقة عند استجابة المفحوص لها، بعدها تتم خطوة التحقيق من الاستجابات التي يعطيها المفحوص والتي تكون بالبطاقة العاشرة ثم التاسعة فالثانية وهكذا. والهدف من ذلك توضيح مكان الاستجابة هل تشمل الشكل ككل آم جزء، ثم معرفة التتابع والترتيب الذي سارت عليه استجابة الفرد، ومعرفة اسلون التنظيم الذي اعتمده المفحوص في جمع الوحدات الصغرى في وحدات اكبر منها كذلك لا بد من معرفة الاستجابات المألوفة وغير المألوفة حيث ترد الاستجابة غير المألوفة أو الأصيلة مرة واحدة في كل (100) تقرير عادي.

وفيما يتعلق بالتصحيح لاختبار رورشاخ تعطى الدرجة بناءاً على المحتوى والموقع والعوامل المحددة ويشمل صنف المحتوى أنواعاً من الاستجابات مثل: البشرية (رجل يرقص)، الحيوان (وطواط)، الجنس (ثدي امرأة)، الطعام(قطعة مقلية من اللحم)، إذا كانت معظم إجابات المفحوص حيوانية يؤخذ ذلك كمؤشر على الذكاء الواطيء، وإذا كانت استجابته في صنف معين أكثر من بقية الأصناف فإن ذلك دليل على وجود حاجات شخصيته تتعلق بذلك الصنف.

إن الموقع يشير إلى جزء من بقعة الحبر التي تستثير استجابة خاصة وتأخذ الاستجابة درجة (W) إذا استخدمت كل البقعة في الاستجابة مثل ((البقعة تمثل خفاشاً))، كما أن الاستجابة تعطى في التصحيح الرمز (D) إذا كان الموقع يدرك كأجزاء فرعية من البقعة، والإشارة إلى جزء كبير من البقعة الحبرية على أنها (رأس طير) مثلاً تعتبر استجابة نموذجية للنوع (D) ويعطى الرمز (Dd) في التصحيح إذا كان الموقع صغير يمثل جزءاً غير ملحوظ بصورة مباشرة. أما بالنسبة لتصحيح العوامل المحددة فهو أكثر تعقيداً من تصحيح المحتوى والموقع. إن الاستجابات من نوع الشكل يعتقد بأنها تشير إلى أسلوب واقعي عقلاني في التعامل مع الحياة، أما استجابات اللون فيعتقد بأنها تشير إلى إرجاع انفعالية، وإذا أعطى شخص استجابات شكلية أكثر من اللونية بصورة ملحوظة، يؤخذ ذلك كمؤشر على الافتقار الانفعالي، أما إذا ساد اللون على الشكل فيعتقد بأنه مؤشر على صنف الضبط الانفعالي. وتعتبر الاستجابات الحركية مؤشراً على تمني التصور، وعندما تتغلب استجابات الحركة على اللون فيعني أن الفرد منطوي. أما إذا تغلبت استجابات اللون على الحركة فيقال أن الشخص على درجة عالية من التوتر وانه يهتم بالحياة الخارجية أكثر من الداخلية. إن تفسير نتائج اختبار رورشاخ يتم في ضوء التحليل النفسي وسيكولوجية الأعماق، أن القائم بالاختبار لا يعتمد فقط على ملخص الاستجابات في التفسير وإنما يقوم بتسجيل الملاحظات عن الكثير من العلاقات المعقدة بين المحتوى والموقع والعوامل المحددة للاستجابة.

وهناك أساليب أخرى في الاختبارات الإسقاطية وهي التداعي الطليق وتحليل الأحلام إذ تقوم الطريقة التي أطلق عليها (بروير) اسم الشفاء بالكلام أو التفريغ قوامها أن يروي المريض التفاصيل التي لا يتسنى ظهور أعراضه للمرة الأولى وكان يعقب ذلك اختفاء الإعراض، وقد طور فرويد هذا النهج بالتدريج وخرج منه بطريقته الخاصة به وهي طريقة التداعي الطليق التي يقول عنها ارنست جونز إنه احد عملين عظيمين في حياة فرويد العلمية.. إن جوهر طريقة التداعي الطليق هو

أن يطلب من المريض أن يقول كل شيء يرد على شعوره مهما بدا ذلك سخيفاً أو غير لائق. إن طريقة التداعي الطليق على عكس طريقة التفريغ لا تقف عند منشأ الأعراض، بل أنها تسمح أو بالأحرى تتطلب أن يتحدث المريض عن كل شيء وأي شيء يخطر بباله دون أي قيود ودون أية محاولة لإلباس ذلك ثوباً من المنطق والمعنى المنظم المعقول، ودور المعالج دور سلبي إلى حد كبير إنه يجلس وينصت ويحث المريض أحياناً فيسأل عندما يجف معين الكلام لدى المريض لكنه لا يقاطع المريض أثناء حديثه، ويستلقي المريض على أريكة في حجرة هادئة بغية خفض تأثير المشتتات الخارجية إلى أقصى الحدود.

وقد لاحظ فرويد أن المريض عندما تتحقق هذه الشروط يشرع آنذاك في الحديث عن ذكريات تتناول خبرات الطفولة المبكرة وقد زودت هذه الذكريات فرويد بأول استبصار حقيقي له في تكوين بناء الشخصية وتطورها التالي...وتحليل الأحلام ليس طريقة منفصلة عن طريقة التداعي الطليق فهو نتيجة طبيعية لما يطلبه من المريض في الحديث عن كل ما يخطر بباله وقد تذكر مرض فرويد الأوائل أحلامهم ثم شرعوا في الإدلاء بمستدعياتهم الطليقة لهذه الأحلام وسرعان ما تحقق فرويد من أن هذه الأحلام المذكورة وما يصاحبها من مستدعيات طليقة كانت مصادر ثرية بصفة خاصة لديناميات الشخصية الإنسانية.... ووفقاً لذالك صاغ فرويد نظريته الشهيرة التي تؤشر أن الحلم يعبر عن نشاط المحتويات اشد العقل البشري إمعاناً في البداية. وقد دعى فرويد العملية البدائية التي تخلق الحلم بالعملية الأولية وفي الوقت الذي يعتمد فيه فرويد على تفسير الحلم الواحد يعتمد يونج على تفسير سلسلة من الأحلام معاً، كما استخدم طريقة التخيل الايجابي حيث يطلب من المفحوص فيها آن يركز انتباهه على إحدى صور الحلم المؤثرة ولكنها غامضة مبهمة أو على صورة بصرية تلقائية ويلاحظ ما يحدث للصورة أيجب إيقاف ملكات النقد وملاحظة ما يحدث وتسجيله بموضوعية مطلقة، فإذا ما روعيت هذه الشروط فإنه عادة ما تمر الصورة سلسلة من التغيرات تخرج إلى الغور كتلة من المواد اللاشعورية.

تصنيفات أخرى لقياس الشخصية

1- تصنيف وكنس Wiggins:

وضع وكنس عام 1973 التصنيف التالي للاختبارات الشخصية وهي:

أ- أساليب الملاحظة:

وتتضمن هذه الأساليب خمسة شروط للحصول على البيانات:

1- المواقف ويقصد بها رصد السلوك كما يحدث على الطبيعة.
2- الملاحظون وهم الأشخاص الذين يقومون بملاحظة الظواهر المدروسة (السلوك).
3- الأدوات وهي المقاييس التي نستخدمها.
4- المناسبات ويقصد بها الفترة الزمنية التي تؤخذ منها العينات السلوكية.
5- الخصائص وهي الأبعاد التي يراد قياسها.

ب- تحليل السلوك:

تعتمد المدرسة السلوكية على هذا الأسلوب في دراسة الشخصية حيث ترى أن السلوك هو وظيفة للظروف البيئية التي تنتزع الاستجابات السلوكية

ج- الاختبارات المبنية (المحددة):

إن هذه الطريقة تتيح الفرصة للباحث لتحليل وانتقاء الفقرات الجيدة ذات الثبات والصدق العاليين وقد ثبت أن هذه الطريقة أكثر صدقاً من بقية الأساليب الأخرى ولها عدة فوائد منها مسألة التقنين والتصحيح الموضوعي.

وهناك ثلاثة طرق أساسية لبناء مقاييس الشخصية وهي:

1- المنطقي: وتفترض وجود تطابق بين التقرير الذاتي للفرد وسلوكه الخفي وتفترض أيضاً أن المستجيب صادق وله من البصيرة الذاتية على رؤية (إدراك) السلوك الذي يقوم به.

2- التجريبي: أحسن مثال عليه هو اختبار منيسوتا للشخصية المتعدد الأوجه. وهذه المقاييس تعتمد في بنائها على قدرتها على التميز بين أشخاص يتصفون بصفات معينة مثل المرضى النفسانيين وآخرين لا يتصفون بتلك الصفات. (الاسويا عن طريق التجريب)

3- المفهوم: مثلاً هناك مفهوم مشتق من نظرية ويبني الاختبار بناءاً على ذلك المفهوم.

2- تصنيف وأيتلا Whitla:

يقسم اختبارات الشخصية إلى ثلاثة أصناف رئيسية هي:

1- **الملاحظة العيادية:** تؤكد هذه الطريقة بصورة رئيسية على الديناميات السلوكية (العوامل المحركة الداخلية)أنها توجه نحو الحصول على بيانات عن العوامل المتضمنة للدافعية والتوافق والآليات الدفاعية والصرعات، والبيانات التي تستخدم فيها لأغراض التحليل، وتتضمن الرموز اللفظية (اللغة) والتداعي الحر والأحلام وكل من الاختبارات الموضوعية والاسقاطية والمقابلات المقننة وغير المقننة.

2- **الطريقة التجريبية:** إن طريقة قياس الشخصية بواسطة المعالجة (إجراء تغيير أو التداخل في الموقف) الشخص المفحوص أو البيئة الموجود بها تشتمل على مدى واسع من الأجهزة والمصطلحات والجوانب النظرية.

إن الأجهزة قد تكون اختبارات موضوعية أو إسقاطية أو الاستجابات الفسيولوجية، والتعليمات حول كيفية قيام المبحوث بالسلوك أو التغيرات في المنبهات التي يسيطر عليها الباحث تجريبياً.

3- **الطرق الإحصائية:** أن الطريقة باختصار أسلوب في القياس الموضوعي للشخصية ذو أوجه متعددة،ويقوم بوضع التنبؤات والمواصفات على أساس المعالجة الإحصائية للبيانات المجموعة.

ومن الرواد لهذه الطريقة كاتل الذي يطلق عليها (الطريقة المتعددة الأوجه) لكي يميزها عن الطريقة الأحادية أو الثنائية التي يستخدمها التجريبيون في تطبيقهم للمتغيرات المستقلة أو التابعة. وتعتمد هذه الطريقة بصورة خاصة على التحليل العاملي.

3- تصنيف انستازي Anastasi:

يمكن قياس الشخصية بعدة طرق وتبحث ثلاثة طرق معروفة لدراسة الشخصية في الأمور التالية:

1- ما يقوله الفرد عن نفسه (بيانات التقارير الذاتية)

2- ما يقوله الآخرون عن الفرد (بيانات اجتماعية)

3- ما يقوم به الفرد في موقف معين (تقنية الملاحظة)

وإضافة إلى ذلك فإن هناك أسلوباً آخر لتصنيف اختبارات الشخصية وذلك استناداً إلى أسلوب تركيبها.

فإن الطرق الثلاثة المستخدمة لتركيب اختبارات الشخصية هي:

1- criterion keying وتستخدم في قائمة منيسوتا المتعددة الأوجه وقائمة كاليفورنيا النفسية.

2- تحليل العوامل factor analysis (وتستخدم في اختبارات كاتل للشخصية)

3- الأسلوب المنطقي The logical approach (ويستخدم في استبيان بيرنزويز للشخصية)

4- تصنيف كرونباك:

وهذا التصنيف هو مشابه تقريباً لتصنيف انستاري وهو:

1- قياس الشخصية عن طريق التقرير الذاتي.

2- قياس الشخصية عن طريق أحكام الآخرين والملاحظات المنظمة.

3- قياس الشخصية عن طريق اختبارات الأداء حيث يتم فيها وضع الفرد في موقف معين وملاحظة أداءه فيه اثناء قيامها بالمهمات المطلوبة.

5- تصنيف بيرت:

يوضح بيرت ثلاثة مجموعات رئيسية للطرق المسيطرة لتقييم الشخصية:

1- المقابلات التي بواسطتها يقيم الشخص وتوضح له درجة ومرتبة على الأسس والانطباعات من قبل الذين يقومون بالمقابلة.

2- الاختبارات.

3- ملاحظة الشخص في حالة قياسيه في الحالة الطبيعية.

وأشار الدكتور عبد الجليل الزوبعي إلى الأصناف التالية في اختبارات الشخصية:

1- ما يذكره الفرد عن نفسه: ويتم ذلك من خلال مقابلته أو إعطائه استفتاء أو اختبارات للشخصية.

2- تقييم الفرد من وجهة نظر الآخرين، ويتم تقييم الفرد في هذا الأسلوب بواسطة استخدام مقاييس التقدير.

3- قياس السلوك الفعلي للفرد: ويتم ذلك من ملاحظة الفرد إثناء قيامه بسلوك معين وتسجيل استجاباته أو تقييمها.

4- قياس الاستجابات الإسقاطية: ويتم تقويم الشخصية في هذا الأسلوب بناء على ما يسقطه الفرد من ميول واتجاهات وقيم عندما يعرض عليه اختبار للبقع الحبرية مثلاً أو عندما يقوم برواية قصص عن اختبار يحتوي على صور غامضة غير محدودة المعالم، وما إلى ذلك من الأساليب الإسقاطية كما ذكرت سابقاً.

المشكلات التي تواجه اختبارات الشخصية وقياسها

قد تعاني اختبارات الشخصية نفس المشكلات التي تعاني منها الاختبارات السيكولوجية بصورة عامة،غير إنه يضاف إلى ذلك وجود مشكلات خاصة بها ولعل من أهم هذه المشكلات التي تواجه قياس الشخصية هي مسألة التزييف Paking في الإجابة والتي تبدو واضحة في اختبارات الشخصية، إضافة إلى مشكلة (التغير) الذي يحدث في السلوك المراد قياسه باختبارات الشخصية، وهذا يؤدي بالنتيجة إلى ضعف ثبات اختبارات الشخصية، وقد يعود سبب ذلك إلى أن الاستجابة على الاختبار تتأثر بالظروف الموقعية بدرجة أعلى من تأثر المجال العقلي بها.

ومن المشكلات الأخرى التي تواجه اختبارات قياس الشخصية هي تعريف الشخصية، حيث أن هناك تعريفات تقريباً متفق عليها في بعض المجالات،في حين كما يبدو أن تعريف الشخصية لم يكن متفق عليه، لأنه يرتبط بالنظريات المتعددة التي تفسر الشخصية والتي لكل منها مفهوم معين للشخصية،وهذا بطبيعة الحال ينعكس على كيفية قياس الشخصية وما يراد قياسه.

وفي بعض الأحيان قد تظهر بعض الأغراض الأخرى التي ليست من الأغراض المحددة التي بني من اجلها الاختبار وفق نظرية معينة، مما لا يمكن استخدامه لأغراض أخرى إلا إذا وجدنا صدقه مرة أخرى. وهناك مشكلة أخرى تكاد أن تكون من أصعب مشكلات قياس الشخصية، وهي تتعلق بعينات السلوك المقاس، حيث عادة في قياس الشخصية نأخذ عينات من السلوك التي نعتبرها ممثلة للمجتمع الأصلي، في حين قد لا تكون ممثلة وبالتالي سيكون حكمنا أو قياسنا خاطئاً ولا يعبر حقيقة عن السلوك المراد قياسه.

بيد أن هذه المشكلات التي تواجه قياس الشخصية، لا يعني الابتعاد عن قياس الشخصية، وعدم إعداد اختبارات أو مقاييس لها، إلا أنها تفرض على الباحث أن يكون دقيقاً عند إعداد الاختبار، ويضيعها أمامه كي يضع بعض الحلول لتجاوزها أو لتقليل من آثارها، ومن هذه الحلول أو المعالجات هي:

1- أن يبتعد الباحث عن الذاتية في إعداد الاختبار، وعند تفسير النتائج.
2- ينطلق الباحث من طبيعة المفاهيم في المجتمع التي يبنى عليها الاختبار.
3- أن لا يكون في الاختبار مجالاً للغش أو التأثير المباشر لنوع معين من الإجابة.
4- أن يكون مجتمع الدراسة وتفاصيله واضحة أمام الباحث وهذا يرتبط بالتطور النظري له.
5- أن يتم اختيار العينة بشكل دقيق وممثل،ويفضل اخذ عينات متنوعة.
6- اختيار المقاييس التي تقترب من الحياة الواقعية ويبتعد عن الأسئلة المباشرة في القياس.

وفيما يلي نماذج من مقاييس الشخصية:

استخبار ايزنك للشخصية (صيغة الراشدين)

the Eysenk personality inventory

بني الاختبار في عام 1959 من قبل Eysenk H. J. and S. Eysenk وللاختبار صورتان متكافئتان وتتكون القائمة من بعدين أساسيين هما:

1- الانبساط: وهو بعد ثنائي القطب يجمع بين المنبسط في طرف وبين المنطوي من طرف آخر
2- العصاب: وهو بعد ثنائي القطب يجمع بين ذوي الاستعداد للإصابة بالاضطراب العصابي وبين ذوي الدرجات الدنيا على العصاب. وتحتوي كل صورة على 57 فقرة منها 24 فقرة لقياس العصاب و 24 لقياس الانبساط و 9 لقياس الكذب.

يستخدم الاختبار في مجال التشخيص والعلاج في مجال الاضطرابات السلوكية ويطبق الفاحص المقياس فرديا أو جماعيا. قنن المقياس على فئتين إحداهما فئة سوية وأخرى غير سوية وتكونت العينة من 2000 فردا وضمت الفئة السوية أفرادا من

مهن مختلفة كالمهندسين والمعلمين والطلاب والمدراء في حين ضمت الفئات غير السوية أفرادا عصابين وقد استخرجت المتوسطات والانحرافات المعيارية لكل فئة.

واستخرج الثبات بطريقة إعادة الاختبار فتراوحت معاملات الثبات بين (0,84) و(0,95) للأبعاد المختلفة.

وتم التوصل إلى دلالات الصدق تمثلت في قدرة الاختبار على التمييز بين الأفراد الذين تم تصنيفهم في إحدى الفئات التشخيصية وفقا لأحكام الخبراء من الأطباء الاختصاصيين النفسيين.

وتم تكييف الاختبار على البيئة العربية الدكتور احمد عبد الخالق وطبقه على البيئة الكويتية عام 1991 وتكون الاختبار من 90 فقرة يجاب عنها بنعم أو لا مقسمة على أربعة مقاييس فرعية وهي:

- 24 فقرة لقياس الذهانية
- 20 فقرة لقياس الانبساط
- 23 فقرة لقياس العصابية
- 23 فقرة لقياس الكذب

وطبق الاختبار أيضاً على البيئة السعودية والمصرية وحسب الصدق بطريقة التحليل العاملي والصدق التقاربي والاختلافي وفيما يلي نماذج من فقرات الاختبار:

1- مقياس الذهانية

- هل يقلقك أن تكون عليك ديون؟
- هل العادات الحميدة والنظافة لها أهمية كبيرة عندك؟
- هل تستطيع أن تفهم بسهولة مشاعر الآخرين عندما يكلمونك عن مشاكلهم؟
- هل تشعر بالقلق إذا عرفت أن هناك أخطاء في عملك؟

2- مقياس الانبساط

- هل تحب الخروج كثيرا؟

- هل تبادر أنت عادة بتكوين أصدقاء جدد؟
- هل تفضل القراءة أكثر من مقابلة الناس؟
- هل تحب الاختلاط بالناس؟

3- **مقياس العصابية**

- هل يتقلب مزاجك كثيرا؟
- هل تعتبر شخصك عصبيا؟
- هل تعاني من قلة النوم؟
- هل تشعر غالبا بالوحدة؟

4- **مقياس الكذب**

- هل أنت شخص كثير الكلام؟
- هل تتفاخر بنفسك قليلا من حين لآخر؟
- هل حدث مرة أن تأخرت عن موعد أو عمل؟
- هل تؤجل أحيانا عمل اليوم إلى الغد؟

اختبار هولزمان لبقع الحبر Holtzman inkblot technique

بني من فبل Wayne H. holtzman 1961 للفئة العمرية (5) سنوات فما فوق وللاختبار صورتان A&B وهو اختار اسقاطي طور على غرار اختبار الرورشاخ لبقع الحبر لتلافي بعض العيوب السيكومترية فيه ويتكون من (90) بقعة حبر تشكل (45) زوجا متماثلا بحيث تشبه كل بقعة من الزوج البقعة الأخرى من حيث خصائصها كمثير ومن حيث خصائص الاستجابة لها. يستخدم الاختبار بشكل أساسي في التشخيص الإكلينيكي للحالات المرضية المختلفة.

يطبق الاختبار بشكل فردي. وقنن الاختبار على عدد من العينات من الأفراد العاديين من عمر 5 سنوات وأكثر وفي عدد من الفئات المرضية كالفصاميين والجانحين والمدمنين على الكحول والأطفال المضطربين انفعاليا وغيرهم وقد بلغ

عدد الأفراد في مجموعات التقنين المختلفة حوالي (2000) واشتقت المعايير لهذه الفئات واستخرجت الرتب المئينية.

استبانة الشخصية للمرحلة الثانوية

High school personality questionnaire

بنيت الاستبانة من قبل Cattel & Raymond. B. 1968 وتم مرجعته عام 1984. وللفئة العمرية 12–18 سنة. وللاختبار أربع صور. تتكون كل صورة من صور الاستبانة من (142) فقرة موزعة على (14) بعدا بواقع عشر فقرات لكل بعد إضافة إلى فقرتين لا تنتميان إلى أي بعد وهذه الأبعاد هي:

1- الدفء.
2- الذكاء.
3- الاستقرار الانفعالي.
4- القابلية للإثارة.
5- السيطرة.
6- التفاؤل.
7- الامتثال.
8- الجرأة.
9- الحساسية.
10- الانسحاب.
11- الإدراك.
12- الغرور.
13- ضبط النفس.
14- التوتر.

ويستخدم الاختبار في تقييم الصراعات الانفعالية والاضطرابات السلوكية كما تفيد في مواقف الإرشاد المهني والتربوي وفي أغراض التشخيص والعمل مع الجانحين وفي مواقف عديدة أخرى.

اشتقت فقرات المقياس استنادا إلى نظرية السمات وقد جمعت بعض الفقرات من مقاييس سابقة وانتقيت (560) فقرة وزعت على (14 9 بعدا والفت أربع صور غير متكافئة من المقياس بواقع (142) فقرة في كل صورة تمثل الأبعاد الأربعة عشر وقد وصفت الفرات في الأبعاد باستخدام أسلوب التحليل العاملي.

يطبق الاختبار فرديا أو جماعيا ويستغرق الوقت لتطبيقه ما بين (45-60) دقيقة. استخرج ثبات الاختبار بطريقة إعادة الاختبار ومعاملات التكافؤ بين الأبعاد في الصور المختلفة للاختبار.

توصل إلى دلالات صدق البناء العاملي للمقياس عبر مراجعاته العديدة وقد استخدمت الدراسات المختلفة محكات متنوعة لمعرفة علاقتها بالأداء على المقياس مثل التحصيل والميول الدراسية والتسرب والإبداع. وقد استخدم عدد من المحكات ذات الصفة التشخيصية الإكلينيكية مثل التكيف الأكاديمي والاضطرابات المرتبطة بالقلق واضطرابات الكلام وغيرها.

اختبار هارور لبقع الحبر للتشخيص النفسي

Harrower psychodiagnostic inkblot test

بني من قبل Molly& Molly 1966 يستخدم للأعمار مابين 16-68 سنة. وللاختبار صورة واحدة مكونة من (10) بطاقات تقيس القدرة على تحمل التوتر أو الضغوط النفسية. وهو فعال في مجالات التطبيق الإكلينيكي وتشخيص الشخصية وتقييمها ويستخدم لأغراض الإرشاد والتوجيه وفي عملية الانتقاء والاختيار. طور الاختبار خلال الحرب العالمية الثانية كأحد أجزاء بطارية اختبارات لقياس تحمل التوتر النفسي وهو يعتمد على المسوغات نفسها التي طور من خلالها

اختبار الرورشاخ. يطبق فرديا وجماعيا وفي حالة التطبيق الجمعي يعطى المفحوص التعليمات التي تتمثل في إنه سوف تعرض عليه مجموعة من البطاقات التي تحتوي على بقع من الحبر والمطلوب منه أن يكتب كل شيء يخطر في باله عند رؤيته البطاقة. ويمكن تطبيق المقياس ذاتيا دون حاجة إلى الاختصاصي.

تم بناء معايير الاختبار في عينة مكونة من (485) مفحوصا منهم (315) عاديا و(170) مريضا نفسيا وتراوحت أعمارهم مابين 16-68 سنة ومن كلا الجنسين منهم (243) ذكور و (242) اناث. ويتولى أمر التفسير أخصائي إكلينيكي بطريقة التفسير الخاصة باختبارات الرورشاخ كما يشرحها دليل الاختبار.

اختبار كوميري للشخصية Comery personality scakes

بني من قبل Comery A. L. عام 1970 على طلبة المرحلة الثانوية والجامعة ومن الفئة العمرية (16-60) سنة.

وتعد الأداة متعددة الأبعاد شاملة لتقييم المميزات الأساسية للشخصية ويزودنا الاختبار بثمان درجات تمثل أبعاد الشخصية التالية وهي:

1- الثقة مقابل الميل للدفاع
2- التنظيم مقابل الاضطرار للتنظيم
3- الانسجام الاجتماعي مقابل التمرد
4- النشاط مقابل عدم توافر الطاقة
5- العاطفة المستقرة مقابل العصبية
6- الانبساط مقابل الانطواء
7- الذكورة مقابل الأنوثة
8- التعاطف مقابل التمركز حول الذات (الأنانية)

وكل مقياس من المقاييس السابقة يحتوي على (20) فقرة فضلا عن وجود مقياسين للصدق يحتوي كل منها على (8) فقرات وثانيهما مقياس التحيز للاستجابة

ويحتوي على (12) فقرة وعدد فقرات الاختبار (180) فقرة وزعت بالتساوي على المقياسين وقد صيغ نصفها بشكل ايجابي ونصفها الآخر بشكل سلبي.

يستخدم الاختبار مع الأفراد العاديين للتعرف على مميزات سلوكهم اليومي كما يستخدم لأغراض البحث التي تحاول الربط بين بعض المتغيرات والظواهر والشخصية كما يستخدم لأغراض التشخيص النفسي وفي المجال المهني والصناعي.

قنن الاختبار على عينة مكونة من (362) من الإناث و (365) من الذكور من طلاب الجامعات. واستخرج الثبات عن طريق الاتساق الداخلي باستخدام الطريقة النصفية مصححة بمعادلة سبيرمان براون وتراوحت معاملات الارتباط مابين 0,87–0,96. واستخرجت معاملات الصدق باستخدام التحليل العاملي. واستخرجت دلالات الصدق عن طريق صدق المحك مع عدد من الاختبارات مثل اختبار كلفورد وكاتل وايزنك وبلغ معامل الارتباط بين الاختبار وايزنك 0,76 وأوجدت دلالات صدق البناء للمقياس عن طريق إيجاد معاملات الارتباط بين الدرجات المتحققة على المقياس وعدد من المتغيرات كالجنس وعلاقات الفرد الاجتماعية ونشاطاته المدرسية وهواياته وميوله.

الفصل الثاني عشر

قياس الميول

يحتوي الفصل على:

- مقدمة
- طبيعة الميول
- الميل والمفاهيم الأخرى
- تكوين الميول
- العوامل التي تؤثر في الميول
- تقدير الميول وقياسها
- قائمة سترونك للميول المهنية وقياسها
- بيان المفاضلة المهنية لكودر
- اختبار مسح الميول ليفورد وزيمومن
- اختبار الميول المهنية لأحمد زكي صالح

مقدمة

تعتبر الميول من المتغيرات الهامة في الشخصية، لذا اهتمت بها الدراسات النفسية لكونها ترتبط ارتباطاً وثيقاً بالإقبال على نواحي النشاط في المجالات المختلفة وترتبط كذلك فيما يتجه إليه الأفراد من أنواع النشاط في أوقات فراغهم. والميول من الخصائص التي تؤثر أساساً في تكيف الفرد التربوي والمهني وعلاقاته الشخصية المتفاعلة. والمتعة التي يستمدها من ممارساته المهنية. والجوانب الرئيسية الأخرى خلال تعامله اليومي. والميول تسير جنباً إلى جنب مع الاتجاهات والقيم والدوافع والمتغيرات الأخرى ذات العلاقة لتشكل جانباً هاماً من الشخصية المميزة للفرد. وبرغم أن اختبارات موثوقة وجهت بشكل دقيق نحو قياس واحد أو أكثر من هذه المتغيرات، إلا أن مثل هذه الاختبارات لا يمكن تصنيفها على نحو جامد في فئات منفصلة كميول واتجاهات وقيم، وما شابه ذلك.

والتداخل والتشابك هو القاعدة، أي أن الاستبيان المصمم لتحديد القوة النسبية للقيم المختلفة، مثل القيم العملية والجمالية (الفنية) أو العقلية، قد يشترك في نقاط عديدة مع قوائم الميول، وبالمثل فإن مثل هذا الاستبيان قد يشير إلى قياس اتجاهات الشخص نحو المعرفة الصرفة، والفن من اجل الفن،والتطبيقات العملية، وما شابه ذلك.

ومن المحتمل أن تكون دراسة الميول قد تلقت قوتها الدافعة من خلال الإرشاد التربوي والمهني، وربما ترجع إلى عملية التصنيف والاختيار المهني، لذا فإن الدراسة المستفيضة للميول وكيف تتكون وتنمو وكيفية قياسها أصبحت من الفروع الهامة في ميدان الإرشاد النفسي وبحوث الشخصية، كما أن دراسة ميول الشخص تكتسب أهمية عملية من خلال وجهات نظر العمال وأصحاب العمل. ولابد من الإشارة خلال الحديث عن الميول إلى أن الإنجاز هو محصلة لشيئين هما، الاستعداد والميل، برغم أن هذين المتغيرين تربطهما علاقة إيجابية إلا أن المستوى العالي في أحدهما لا ينطوي بالضرورة على مستوى عال في الثاني. فإن شخص ما قد يمتلك استعداداً

كاف للنجاح في نمط محدد من النشاط سواء كان مهنياً أو تربوياً أو استجمامياً دون اهتمام أو ميل مماثل. أو إنه قد يكون ميالاً إلى عمل ما، إلا إنه يفتقر إلى الاستعدادات اللازمة لذلك العمل. لذا فإن قياس كلا هذين النمطين من المتغيرات يتيح الفرصة للتنبؤ عن الأداء بفاعلية أفضل مما لو كان القياس بالنسبة لواحد منهما فقط.

ولو نظرنا إلى الدول المتقدمة علمياً وحضارياً نجد أنها تهتم بموضوع الميول أكثر من الدول الأخرى النامية أو غير المتقدمة. ففي الأولى نجد أن هناك موجهين وأخصائيين يعملون لمساعدة الأفراد على الملاءمة بين قدراتهم واستعداداتهم من جهة وبين ميولهم من جهة أخرى. سواء كان ذلك من الوجهة التربوية أو المهنية حيث تجري اختبارات لميول الطلبة، وكذلك اختبار قدراتهم واستعداداتهم ومن ثم يختار الموجه نوع الدراسة التي يميل إليها الفرد والتي تساعده فيها قدراته واستعداداته كذلك من جهة الأعمال المهنية نجد أن هناك اختبارات لمعرفة ميول الأشخاص قبل التحاقهم بأي عمل، فإذا ما تلائم الميل مع القدرة لكان هذا من اكبر العوامل الهامة في نجاحه سواء في دراسته أو عمله، وهذا النجاح يرجع إلى الراحة والسرور الذي يجده الفرد من اختياره لعمل أو دراسة يميل إليها وتؤهله قدرته على الاستمرار فيها. وقبل الخوض في طرق قياس الميول واهم المقاييس شائعة الاستخدام لابد من الإشارة ولو بشكل موجز إلى نقاط هامة يمكن أن تؤدي إلى فهم أوسع وإحاطة اكبر بجوانب الموضوع. إذ إنه ينبغي على المتخصص في بناء المقاييس الخاصة بمجال معين، أن يتعرف بشكل دقيق على طبيعة ذلك المجال، والعوامل المؤثرة فيه، والحدود الفاصلة - أن وجدت - بينه وبين المجالات الأخرى ذات العلاقة، ومراحل تكوينه، لكي يكون ذلك المقياس ناشئاً من صلب وطبيعة ذلك المجال وليس من واد بعيد عنه أو مجاور له. والميل لا يختلف عن أي من هذه المجالات.

طبيعة الميول:

عندما يقوم الفرد بأوجه نشاطه المختلفة، وكذلك عند تعامله مع ما يحيط به، نجد إنه يفضل بعضهما ويرغب فيها ويرتاح إليها، كما نجده يهمل بعضاً آخر وينفر منه فإننا نقول أن هذا الفرد يميل إلى ما يفضله ويرتاح إليه. فهو يجد في ذلك إشباعاً لحاجاته، كما إنه يشعر بحرية فيما يفعله، بينما نقول على البعض الآخر إنه لا يميل إليه. لذا يمكن أن نعرف الميل فنقول "إنه استعداد لدى الفرد يدعوه إلى الانتباه والاستمرار في نشاط ما يثير شيء في نفسه، أو هو القوة التي تدفعه إلى التفضيل بين أوجه النشاط المختلفة" وقد تعددت تعاريف علماء النفس بهذا الصدد ومنها:

1- تعريف كيلفورد: إنه نزعة سلوكية عامة لدى الفرد للانجذاب نحو نوع معين من الأنشطة.

ويعني بقوله (نزعة سلوكية عامة) إنه ليس شيئاً أكثر من كونه سمة عامة. كما يعني (بالانجذاب نحوه) أن الفرد يهتم بـ أو يتجه نحو، أو يبحث عن، أو يهدف إلى الحصول على شيء له قيمة كامنة بالنسبة له، وكون الفرد ينجذب نحو أنشطة معينة معناه أن الميل اقرب إلى أن يحدد ما يفعله الفرد، ا كثر مما يحدد كيف يفعله،وتعريف الميل على هذا النحو يضعه في المجال العام للدوافع فالميول - كالحاجات والاتجاهات - تكون نوعاً من السمات التي يمكن أن يطلق عليها دينامية أو دافعيه.

2- والعالم ويبسترWebster يعرفه بأنه استعداد لدى الفرد يدعوه إلى الانتباه إلى أشياء معينة تستثير وجدانه ومن هذا التعريف نلاحظ إنه يتضمن ثلاث نقاط يجب أن تتوافر وهي:

أ- الانتباه.

ب- الأشياء التي ينتبه إليها الفرد.

ج- الإثارة الوجدانية نتيجة هذا الميل.

3- أما سترونج Strong فيفضل استمرار استعمال الانتباه بدلاً من مجرد الانتباه لان الانتباه يتحول من موضوع إلى آخر بينما يكون الانتباه في الميول مستمراً،لان الميل إلى نشاط معين يجعل الفرد يكرر هذا النشاط ومن هذا نجد أن مبدأ الاستمرار يلحق بالميول أكثر.

4- ميرفي Murphy يرى أن الميل هو اتجاه من شأنه يجعل الإنسان ينتبه لأشياء معينة وما يصاحب هذا الانتباه من شعور.

5- ويعرفه بنجهام Bingham على إنه استعداد من جانب الإنسان لان يستغرق في نشاط ما وهذا الاستعداد ديناميكي،بمعنى إنه يؤثر في سلوك الإنسان.

6- أما الدكتور احمد زكي صالح فيعرفه،بأنه استجابة مقبولة إزاء موضوع خارجي معين وهذه الاستجابة متصلة ومكتسبة، لأن الفرد يتصل بموضوعات العالم الخارجي عن طريق نشاطه في هذه الموضوعات، فإذا ترتب عن هذا النشاط ألوان من الراحة والسرور والمتعة فإنه نتيجة لذلك ينشأ الميل نحو هذا النشاط.

وبرغم تعدد وجهات النظر حول تعريف الميل، إلا إننا يمكن أن نشير إليه على إنه استجابة قبول إزاء موضوع خارجي ولذلك فالميول متعلقة ومكتسبة من تفاعل الإنسان مع بيئته الخارجية، سواء كانت هذه البيئة مادية أم اجتماعية. أي أن الميول تعتبر استعدادات مكتسبة بطريقة معينة لنواحي خاصة من البيئة المحيطة بالفرد.

وهكذا تعتبر الميول نموذجاً سلوكياً من نماذج الدوافع المكتسبة حيث أنها توجه صاحبها للاستجابة بطريقة خاصة لناحية معينة من البيئة التي تحيط به، بمعنى آخر يوجه الميل صاحبه نحو مجموعة معينة من أساليب النشاط المتميزة عن غيرها من الأساليب الأخرى في البيئة أو المجال المحيط به، ومن ثم يشترط الميل الفرد أن يستجيب بطريقة خاصة لموضوع ما أو لمجموعة من الموضوعات في البيئة الخارجية سواء كانت موضوعاً مادياً، أم شخصية معنوية، أو أسلوب نشاط معين.

وطبقاً لتعريف الميل بوجه عام، فإن الميل المهني هو المجموع الكلي لاستجابات

القبول التي تتعلق بمهنة ما. ويتضمن ذلك أنماط الاستجابات الانفعالية،أو العادات السائدة عند الفرد،ومدى ثباته الانفعالي، والصفات المزاجية التي تميزه كالانطواء والانبساط وصفاته الشخصية والاجتماعية، وهكذا يكون الميل المهني ليس وحدة سلوكية أو اختيار واحد، إنما هو تنظيم سلوكي معقد يتعلق بمجموع استجابات القبول إزاء نشاط مهني معين. وقد يتعلق الميل بما يود الفرد أن يمارسه في وقت فراغه فيتحول آنذاك إلى ميل لا مهني مثل الميل إلى مشاهدة السينما أو مشاهدة مباريات كرة القدم وغير ذلك. وهذه الميول قد تتفق في جزء منها مع عمله،أو قد لا يكون بينهما وبين عمله المهني علاقة.

ويظهر أيضاً أن ما يميز الميل هو إنه أحادى البعد أي إنه ذو ناحية واحدة وهي الناحية الإيجابية - فنحن لا نميل إلا للأشياء التي تجلب لنا السرور والراحة - كما إننا قد نميل لأشياء نستطيع أن نعملها بإتقان كالميل للقراءة والنظافة بينما نجد أن هناك ميولاً لدينا ولكننا لا نستطيع أن نؤديها بإتقان فمثلاً يميل شخصاً ما إلى سماع الموسيقى ولكنه لا يستطيع أن يعزف على الآلات الموسيقية كما أن هناك ميولاً سائدة بين الأفراد كالميل نحو الرقي المهني أو الميل نحو الوصول إلى إعجاب الآخرين، وهي ميول سائدة لأنها تؤثر في سلوك الفرد من جميع النواحي.

الميل والمفاهيم الأخرى

لا بد لنا ونحن نتعرض لمفهوم (الميل) من القيام بتحديده بدقة،ولا يتم لنا ذلك من خلال الاطلاع على التعاريف السابقة فقط، وإنما من خلال تميزه عن بعض المفاهيم الأخرى التي قد تكون مقاربة له أو ذات صلة به،تجنباً لحدوث أي التباس أو غموض لدى القارئ ومن هذه المفاهيم (الاتجاه) و (العاصفة) و(الحاجة) و (حب الاستطلاع).

الميل والاتجاه

يختلف الميل عن الاتجاه، حيث أن الاتجاه يتميز بالإيجابية أو السلبية. فاتجاه الفرد إزاء تنظيم النسل قد يكون موجباً أو سالباً، وقد يكون محايداً وهكذا في سائر موضوعات الاتجاهات. أما الميل فإنه اتجاه موجب عادة إذ ليس له إلا ناحية واحدة فقط هي ناحية الإيجاب. فنحن لا نميل إلا للاشياء أو الموضوعات التي نرغب فيها،ولا نحب إلا ما نميل إليه. ولو قرأنا تعريف (البورت) للاتجاه أو تعريف (جيلفورد) أو غيرها من التعاريف لوجدنا بأن الاتجاه- من الناحية السيكولوجية- يتضمن المعتقدات مثلما يتضمن المشاعر وفي هذا تمييز آخر للاتجاه عن الميل.

الميل والعاطفة

من الواضح بأن الميول تشبه العواطف، ولكن الميول تتصل بالنواحي النزوعية، وبما يقوم به الشخص من تصرفات تدل على وجود الميل عنده،أما العاطفة فتغلب فيها الصفة الوجدانية.

الميل والحاجة

الواقع أن إشباع حاجات الفرد الأساسية تجعله يشعر بارتياح وسرور، وكل فرد يميل دائماً إلى إشباع حاجاته. ومن هذا نجد أن هناك علاقة بين الميول والحاجات، أي أن إشباع كل الحاجات الأساسية للفرد يكون أساساً لميوله أي إشباع حاجاته بطريقته الخاصة.

الميل وحب الاستطلاع

هناك فرق بين الميل وحب الاستطلاع، فنجد أن حب الاستطلاع نزعة طارئة قد يحيد عنها الفرد إذا لم يجد ما يشبع حاجته النفسية أي إنه لا يستمر فيه، بينما تلازم الميل صفة الاستمرار. كما نجد أن الشخص الذي يدفعه حب الاستطلاع إلى شيء ما ثم يجد في هذا الشيء ما يسبب له الشعور بالراحة والسرور فإنه يتولد هنا

ميل تجاه هذا الشيء. أي أن حب الاستطلاع يميل ميلاً طارئاً بينما يميل الميل ميلاً حقيقياً.

كيف تتكون الميول

الميول تتكون بالتدريج وتنمو مع الزمن، وتتأثر بالعوامل الخارجية التي قد تقويها أو تضعفها وخلال المراحل التي تنمو فيها الميول يمكن التمييز بين نوعين منها وهي:

1- الميول التي تبنى على أسس غريزية أو انفعالية أو عاطفية أو حاجات نفسية. وهذه تكون أكثر ثبوتاً ونمواً، كالميل المتصل بعاطفة جنسية، أو الميل نحو الأمور الاجتماعية، وذلك الميل المبني على الحاجة إلى الشعور بالأمن أو الانتماء للجماعة، أو الميل نحو الهوايات التي يغلب فيها عنصر اللعب.

2- الميول الثانوية أو المشتقة كتلك التي تتصل بالقيام بعمل أو مهنة حيث لا يظهر الميل للإقبال على مثل هذا النوع من النشاط إلا بسبب شعور الشخص بضرورة هذا العمل، ولكن مثل هذه الميول البعيدة الصلة بالحاجات والدوافع النفسية تكون أقل ثبوتاً وأسهل تحولاً وأكثر قابلية للتغيير.

ويمكن القول بأن النوع الأول من الميول تكون فيه صفة التلقائية، بينما النوع الثاني يكون مرتبطاً بأهداف وأغراض خارجية يعمل الشخص على تحقيقها مضطراً أو راغباً.

والطفل في سنواته الأولى يكاد يتركز نشاطه واهتمامه في نفسه، ونحو الطعام والحماية ونحو الوالدين والجو المنزلي، ثم تتحول ميولهم عندما يكبرون قليلاً بحيث يغلب فيها نواحي النشاط وكسب المهارات والتعلم،وعموماً فإن ميول الأطفال لا تبدأ في الظهور أو الوضوح إلا فيما بعد السنة الثامنة تقريباً. وهنا يبدأ اهتمام الطفل بالعالم الخارجي المحيط به، فنجده يميل لنوع خاص من المهنة كالطيران مثلاً أو الطب أو الهندسة..... الخ.

والميول من هذا النوع لا تلبث أن تختفي مع الوقت، ويظهر في أواخر مرحلة الطفولة ميل عام في معظم الأطفال يرتبط بتركيب الأشياء وصنعها. ويرتبط ظهور هذا الميل بنمو القدرة العضلية الدقيقة ويستطيع البنين في تلك المرحلة استخدام الآلات النجارة والأشغال اليدوية بصفة عامة، كما تستطيع الفتيات استعمال أدوات الحياكة وأعمال الإبرة والتريكو. وقد يصل هذا الميل لدرجة كبيرة من الشيوع في الأطفال، حتى يظهر وكأنه حاجة أو دافع ويصاحب بانفعال المرح والسرور للإحساس الطفل في هذه المرحلة بأنه صائغ ماهر استطاع أن يبتكر ويصنع وقد يستفيد المدرس من فهمه لهذا الميل بإيجاد المجال العملي التطبيقي للأطفال في المدارس.

أما في دور المراهقة فنجد فترة يحدث فيها تغيير كبير في الميول،حيث يحدث التحول نحو الأمور الدينية والاجتماعية والخلقية والجمالية. وفي هذه الفترة تتضح عنده ميول لها أهمية كبرى من وجهة نظر علم النفس،وهي الميول الخاصة التي تعبر عن استعداد الشخص الطبيعي للنمو والبروز في مجال هذا الميل. وهي ميول تؤدي غالباً للكشف عن قدرة خاصة أو موهبة معينة. ومن هنا ظهرت الحاجة الملحة لعملية كشف هذه الميول في مرحلة المراهقة كي يمكن عن طريق معرفتها تهيئة الجو المناسب لعملية التوجيه التعليمي المهني بعد ذلك.

وفي دور الرجولة يتجه الميل والاهتمام نحو تكوين العائلة والعمل على الكسب والإنتاج. أما في الكبر فتتجه الميول نحو الأعمال المتصفة بالهدوء والبعد عن المنافسة كارتياد المجتمعات المسلية، وتنسيق الحدائق، والمساعدة في الخدمة الاجتماعية.

ويتضح من خلال الأبحاث والدراسات الطويلة والمستعرضة- التي أجراها (سترونج) الذي كرس حياته كما وجه تفكيره وطاقته وجهد معظم طلابه إلى العمل في هذا المجال وبخاصة دراسته التي أجراها على عينة من الرجال بلغ حجمها (2340) رجلاً تراوحت أعمارهم بين (20-60) سنة وقام بتقسيمهم إلى أربع مجموعات متوسط أعمارهم (55,45,35,25) وقد وجد بأن بدء ظهور نماذج الميول

حوالي سن (15) وتبلورها حوالي سن (18) وثبوتها بين سن (18-21). ويستنتج سترونج ذلك من المتغيرات الحاصلة في الميل عبر الأعمار من سن (15) وحتى سن (25) فالتغير الحاصل في الميل للسنة الأولى من سن (15,5-16,5) يقدر بثلث التغير، والتغير الحاصل في الميل للسنتين اللاحقتين أي من سنة (16,5-18,5) يقدر بثلث التغير الثاني ويحدث ثلث التغير الأخير في السنوات السبع الباقية. كما وجد بأن تغير ميول الرجال بين سن (25-55) هو تغير ضئيل جداً وان مثل هذه التغيرات إذا حصلت فليس لها أية دلالة. وهذا يعني أن الفرد في سن (25) يكون إلى حد كبير ما سيصبح عليه بعد ذلك، بل إنه يمكننا القول أن الفرد في سن العشرين يكون قد أكتسب أوفر قسط من ميوله التي يمكن أن يكتسبها طيلة حياته. ويحاول بعض العلماء ومنهم (سولينبرجر Sollenberger) تفسير ظاهرة تغير الميل الواضح بين سن (15-25) وتغيره الطفيف بعد ذلك تفسيراً فسيولوجياً يرجعه إلى النشاط الهرموني المتزايد للأولاد أبان مرحلة المراهقة والمراهقة المتأخرة، والذي يأخذ بالنقص التدريجي بعد سن (25) وقد أسفرت دراسة سولينبرجر عن كشف العلاقة بين درجات الطلاب في اختبار سترونج للميول المهنية وبين كمية الإفرازات الهرمونية لديهم.

العوامل التي تؤثر في الميول

هنالك جملة عوامل تؤثر في الميول منها:

1- الوراثة:

حيث ظهر من الأبحاث التي أجراها علماء النفس أن الوراثة لها تأثير في الميول، وهذا التأثير يظهر واضحاً في حالة التوائم المتشابهة أكثر من التوائم غير المتشابهة في الجنس، كما إنه يقل في حالة الإباء والأبناء بدرجة اكبر من حالة التوائم.

2- البيئة:

حيث أن الميول مكتسبة ويتعلمها الفرد،فنجد أنها نتيجة طبيعية لتفاعل الفرد مع البيئة المحيطة أي نتيجة للخبرات التي يمر بها الفرد، فالميول المختلفة للأفراد تكون نتيجة اختلاف هؤلاء من حيث نشأة كل منهم في بيئة تختلف عن الأخرى، حضرية أو ريفية، متقدمة أو متأخرة، فقيرة أو غنية.

أن الظروف البيئية المناسبة تسهم في تكوين الميل وتنشطه وترعاه وتنميه. وذلك مثل تلك الميول المرتبطة بثقافة الوالدين، حيث يلاحظ في كثير من الأحيان أن ابن الموسيقي عادة ما يلعب الموسيقى لأنه يعيش في هذا الجو، ولديه الآلات والتوجيهات المناسبة، ويتدرب مبكراً عليها، ويجد باستمرار التشجيع والتوجيه وبذا يتكون لديه الميل، كما وان الطبيب أو المهندس أو المحامي غالباً ما يؤثر في توجيه النمو العقلي لأبنائه كي يسلكوا نفس طريق الآباء التعليمي. وتنشأ الميول في هذه الحالة من خلال تقليد الآباء خصوصاً إذا ما اعتبر الابن أباه مثلاً يحتذي به.

3- المستوى الاجتماعي:

هناك فروق بين ميول الأفراد الذين ينتمون إلى المستويات الاجتماعية المختلفة، فقد لاحظ سترونج أن الأفراد الذين يعملون في المهن الراقية، سواء كانت فنية أو إدارية يختلفون في أنماط ميولهم عن الأفراد الذين يعملون في المهن الدنيا، فالأولون يميلون إلى أوجه النشاط الأدبي والقانوني وهم اقل ميلاً إلى الخدمة الاجتماعية والميول المتعلقة بالنواحي التطبيقية. وقد لوحظ أن ذوي المهن الراقية يحصلون على درجات عالية في اختبار مستوى الميول المهنية على الرغم من اختلاف مهنهم وبالتالي على الرغم من ميولهم المهنية.

4- السن:

السن له تأثير في تبلور الميول ونضجها وهذا ما تمت الإشارة إليه عند التعرض لموضوع كيفية تكون الميول.

5- الدين والعادات والتقاليد:

هذه العوامل ذات تأثير كبير على ميول الأفراد، فنجد أن أفراد البيئات المتقدمة حيث تقل التقاليد والمعتقدات البالية يميل أفرادها إلى التمتع بالحياة بجميع أشكالها، وفي شتى صورها حيث إنه لا يوجد ما يقيد ذلك. بينما البيئات ذات التقاليد المتزمتة يميل أفرادها إلى تصرفات وأشياء معينة. وهذا كله يتأثر أيضاً بالدين والمعتقدات السائدة في كل بيئة.

6- الجنس:

من المسلم به أن هناك فروقاً في الميول بين الذكور والإناث. فقد لوحظ في اختبارات الميول بصورة عامة أن الفروق بين الجنسين تتلخص في أن الإناث كن أميل إلى أوجه النشاط المتعلقة بالمنزل والحياة الأسرية والى الاهتمام بنواحي الجمال والزينة والألوان، أي يميلون إلى الأشياء المتعلقة بالفنون أكثر من غيرها. في حين أن الرجال كانوا يميلون إلى أوجه النشاط العلمي والميكانيكي، والنشاط المتعلق بالأعمال التجارية والاقتصادية كما أنهم كانوا يميلون أيضاً إلى المغامرة والإثارة أكثر من البنات كما وجد بأن الرجال يفضلون أوجه النشاط المتعلقة بالعدد والآلات والرحلات والنشاط الجسمي العنيف والى المخاطرة ويميلون إلى العسكرية، والى أنواع الترفيه التي تحتاج إلى مجهود عضلي كبير. كما وجد أن الإناث يملن إلى الانطواء أكثر من ميل الرجال. وكذلك تميل الإناث إلى الأعمال الكتابية والتدريس والخدمة الاجتماعية والمهن الأدبية، ويملن إلى بعض أنواع الترفيه مثل الزيارات والحفلات ويملن إلى الروايات والقصص الأدبية والغرامية.

7- المهن:

وجد سترونج من أبحاثه الطويلة أن هناك تشابهاً في الميول بين الذين يعملون في مهنة يتفقون في ميولهم، وهذا الاتفاق يميزهم عن أفراد المهن الأخرى.

تقدير الميول وقياسها

طبقاً لوجهة نظر (سوبر Super) في تعريفه للميل، فإنه يشير إلى وجود أربعة تفسيرات لكلمة الميل. ويتصل كل تفسير بطريقة معينة من الطرق التي نتعرف بها على الميول. فكان معنى الميل يتوقف على الطريقة التي تتبع في الكشف عن الميول فهناك:

1- الميول المعبر عنها لغويا:

حيث يعبر الفرد عن ميله أو عدم ميله لشيء معين بمجرد القول، بأنه يحب هذا أو لا يحب ذاك، وتقاس هذه الميول بطريقة الاستفتاءات. وقد بينت البحوث أن مثل هذه الميول التي يعبر عنها المراهقون بهذا الشكل ليست ثابتة ومتغيرة ولا يمكن التنبؤ منها بميولهم في المستقبل ويتوقف مدى ثبات مثل هذه الميول على مدى ما بلغه الفرد من نضج.

الميول الظاهرة:

وهي الميول التي تتضح عن طريق أنواع النشاط أو العمل التي يقوم بها الفرد في حياته اليومية. فإذا كان لدينا تلميذ من التلاميذ يصرف نقوده في هواية التصوير مثلاً، ويقضي وقتاً طويلاً في هذه الهواية. فهذا نشاط يدل على ميله فعلاً للتصوير.

ويجب الاهتمام هنا بالميل الذي يظهر نتيجة للنشاط ذاته، وليس نتيجة للاهتمام المصاحب لهذا النشاط، فمثلاً الطالب الذي يقوم بالتمثيل في مدرسته قد يكون هدفه البحث عن صداقات وعلاقات جديدة وليس هو الميل للتمثيل. لهذا السبب لا يستخدم الميل الظاهر كوسيلة لمعرفة نوع الميل في كثير من الدراسات. كما أن هناك سبباً آخر وهو البيئة والظروف الاجتماعية والاقتصادية قد تحدد فرص إظهار الميل وهذا يظهر في الشخص الذي يكون لديه ميل نحو التصوير ولكن ظروفه وظروف أسرته المادية تمنعه من إظهار هذا الميل والاستمرار فيه.

3- الميول التي تقيسها الاختبارات الموضوعية:

ومن هذه الاختبارات يتضح ميل الفرد تجاه موضوع معين ويغلب على هذا النوع من الاختبارات صفة الثبات، فنجد أن الشخص الذي يميل إلى رياضة معينة مثل لعبة التنس، فإنه قد يعطينا نتيجة لهذه الاختبارات كل ما يتعلق بها من أطوال اللعب وطريقة اللعب، وتاريخ نشأتها وقوانينها والدول المتقدمة فيها بل قد يصل ميله إليها في إنه قد يلم بالحياة الخاصة لأبطالها كظروفهم الاجتماعية ونشأتهم.

4- الميول التي تقيسها الاستفتاءات (الميول الحصرية):

ونتعرف عليها بطريقة مشابهة لطريقة الاستفتاءات التي تنبع في الكشف عن الميول المعبر عنها لغوياً كما في النوع الأول، غير أن الطريقة هنا تختلف عن طريقة الاستفتاءات في أن كل سؤال في القائمة التي تختبر الميول يكون لنوع الإجابة عليه درجة معينة، ولا تكون درجة الفرد على الاختبار هي مجموع درجاته على المفردات كما هو الحال في الاستفتاءات. وخير مثال لهذه الطريقة اختبارات سترونك وكودر، وعلى أية حال فإن سوبر يلخص الاتجاهات الحالية في عمل اختبارات الميول وهي:

أ- استعمال مفردات سهلة ومألوفة لمن يطبق عليهم الاختبار حتى لا يعوق ذلك إظهار ميولهم.

ب- استعمال الاختبارات الموضوعية التي تقيس المعلومات.

ج- محاولة قياس عوامل محددة على أساس التحليل العاملي بدلاً من قياس ميول عامة لا تمت بصلة إلى المهن المختلفة.

وعموماً فإنه يبدو بأن أكثر الطرق صراحة وملائمة لتحديد ميول شخص ما في أنماط العمل المختلفة،والمناهج التربوية والنشاطات الترويحية،هي-ببساطة- في توجيه السؤال إليه. إلا أن الباحثين الأوائل اكتشفوا منذ وقت مبكر بأن الإجابة على الأسئلة المباشرة عن الميول،غالباً ما تكون غير حقيقية وسطحية وغير واقعية

كما أنها لا تمثل الميل الحقيقي للفرد. إن معرفة أسباب هذه الحالة ليست عملية صعبة ذلك أن تقييم الميول الشخصية لفرد ما يتطلب استبصار وخبرة ملحوظة قد لا تتوفر لديه، فكثير من الناس ليست لديهم المعرفة الكافية بالمهن أو الموضوعات أو الأنشطة المختلفة التي يريدون القيام بها، ومن ثم ليست القدرة على الحكم ما إذا كانوا يحبون أو لا يحبون بالفعل كل ما يتضمنه اختيارهم حالياً. إن ميلهم أو افتقارهم إلى الميل- للعمل قد ينشأ عن فكرة قاصرة ذات أفق ضيق من التفكير عما يستلزمه العمل في ذلك الميدان يوماً بعد يوم. هذا من جهة، ومن جهة أخرى فإن تفشي صور وأفكار عقلية فيما يتعلق ببعض المهن تجعل الفرد يقيم تلك المهن بغير ما هي عليه في الواقع، وقد يكون ذلك التقييم اقل أو أكثر مما عليه في الواقع. فحياة الطبيب الاعتيادي أو المحامي أو المهندس مختلفة تماماً عن الروايات المعروضة في دور السينما أو التلفزيون والمجلات المتداولة. لذا فإن المشكلة هي أن الأشخاص نادراً ما يكونوا في وضع يمكنهم من معرفة ميولهم الخاصة في الميادين المتعددة قبل مشاركتهم الفعلية في تلك الميادين، وبمرور الزمن فأنهم سيدركون فائدة مثل هذا الاحتكاك الشخصي، إلا أن الانتفاع من التجربة قد يكون متأخراً، وربما يكون التغيير آنذاك مدمراً جداً.

لهذا السبب فإن هناك إدراكاً مبكراً بين المهتمين بقياس الميول،بضرورة استكشاف الوسائل غير المباشرة والبارعة في تحديد الميول. إن واحدة في أكثر هذه الوسائل نجاحاً هو ما تمخض عن السيمنار الجامعي حول الميول والذي أقيم في معهد كارنيجي التكنولوجي خلال العام الدراسي(1919-1920) حيث أعدت فيما بعد اختبارات مقننة متعددة للميول كنتيجة للعمل الذي بدأه معدوا هذه الاختبارات عند مشاركتهم في السيمنار المذكور. إلا أن المقياس الذي حقق نجاحاً متميزاً هو اختبار الميل المهني لسترونك وقد اخضع هذا الاختبار للبحث المتواصل من حيث التنقيح والتوسيع، وهو يختلف عن جميع الاختبارات المبكرة الأخرى.

أن قوائم الميول المطورة من قبل مجموعة (كارنيجي) وضعت موضع الاستعمال

اثنان من الإجراءات المبتكرة الرئيسية. الأول: يتضمن الفقرات التي تتعامل مع رغبة أو عدم رغبة الأشخاص في تشكيلة واسعة من النشاطات المحددة والأهداف أو أنواع الأشخاص الذين يقابلهم عادة في حياته اليومية. والثاني: تكون فيه الإجابات مكيفة تجريبياً بالنسبة للمهن المختلفة. وقوائم الميول هذه كانت من بين الاختبارات الأولى التي تستخدم مفاتيح محكية للفقرات. والتي اتبعت فيما بعد في تطوير قوائم الشخصية كال (MMPI) والـ (CPI). لقد وجد بأن الأشخاص المنشغلين بمهن مختلفة يتصفون بميول عامة تميزهم عن الأشخاص في المهن الأخرى. والحقيقة أن هذه الاختلافات في الميول لا تعود فقط إلى المواضيع المرتبطة مباشرة بنشاطات العمل وإنما أيضاً إلى العمل المدرسي والهوايات والألعاب الرياضية وأنواع الألعاب الأخرى والكتب التي يستخدمها والعلاقات الاجتماعية، وكثيراً من المظاهر الأخرى في الحياة اليومية. لذا فإنه من المعقول أن نسال الشخص حول ميوله عن الأشياء المألوفة نسبياً،وبذلك يمكن تحديد إلى أية درجة تكون ميوله مشابهة لميول الأشخاص المشتغلين بنجاح في المهن المختلفة.

وإضافة إلى ما ذكر فهناك نواحي عديدة تساعد في الكشف عن الميول وقوتها لدى الفرد بالنسبة لموضوع من الموضوعات، منها:

1- **المعلومات أو المعرفة:** من الممكن أن يتخذ مدى معرفة فرد ما بموضوع من الموضوعات كمقياس لميله نحو هذا الموضوع. فلو فرضنا أن لدينا عدداً من الموضوعات وكان الفرد أكثر الماماً بأحدها. كان ذلك بمثابة دليل على ميله لهذا الموضوع.

2- **التداعي الحر أو المقيد:** عند استجابة فرد ما لكلمات مختارة اختياراً جيداً فإن طبيعة وتكرار استجابات التداعي قد يكشف عن نمط الميل السائد لدى الفرد.

3- **التفضيل:** الحب والكراهية قد يطلب من الشخص ترتيب قوائم فقرات أو وحدات تعرض عليه بحيث لا يكون القصد من ذلك واضحاً للمفحوص.

وقد تتضمن هذه صوراً من الترويح أو الواجبات الخاصة أو الموضوعات المدرسية أو الكتب وما شابه ذلك،وتتخذ استجابات الفرد المعبر عنها وترتيبه لها وتفضيله لبعضها كمقياس للميل.

4- **الجدول الزمني:** تسجيل توزيع كامل للفترات الزمنية التي يكرسها الفرد لنواحي النشاط المختلفة وأشكال الترويح مما يكشف عن أنماط ميل الفرد. وقد يكون هذا النوع من الأدلة من أنسب الأشياء التي يمكن القيام بها لدراسة الميل عند صغار الأطفال.

5- **ملاحظة السلوك:** يلاحظ سلوك الفرد بدقة وكذلك أنواع قراءاته واتصالاته بالآخرين وسرعة التشتت والانتقال من عمل لآخر.

وقد يلجأ الباحث عند تقييم ميل فرد ما إلى أكثر من طريقة. فقد يتخذ الباحث مثلاً الملاحظة والجداول الزمنية معاً كوسيلة لتقييم ميول الأطفال كما قد تستغل اختبارات إكمال الجمل وموازين التقدير في معرفة ميول الأفراد.

إن معظم مقاييس الميول تعتمد على اللغة في صياغة الأسئلة. ولن يتمكن الفرد من الإجابة عليها إلا إذا تمكن من فحصها. وللتغلب على الصعوبة اللغوية في هذه المقاييس كانت هناك محاولات لاستغلال الصور في اكتشاف الميول، إذ ابتدع العالم(سوبر) من جامعة كلارك وتلاميذه طريقة تعرض فيها صور مختلفة بالفانوس السحري وتعطي هذه الصور معلومات عن نواحي النشاط المختلفة في مهنة من المهن. وهذه الطريقة طريقة موضوعية لمعرفة الميول لقياس المعلومات.

وهناك اختبار (لامونز وآخرون) عبارة عن عشر لوحات من الصور لقياس ميول النساء وثمانية للرجال، وكل صورة تمثل حرفة من الحرف أو مهنة من المهن للتعرف على معلومات الأفراد وميولهم للمهن المختلفة. وهذا الاختبار مماثل للاختبار السابق.

وعلى أية حال فإننا سنركز في عرض بعض الاختبارات المعيارية النموذجية

المصممة لقياس الميول والتي تعبر عن الاختبارات الحصرية - سابقة الذكر- وهي قائمة سترونك للميول المهنية، وقوائم الميول لكودر.

وهذه الاختبارات هي من نوع قوائم التقرير الذاتي،أي اختبارات الورقة والقلم إذا أن غالبية مقاييس الميول هي من هذا النوع،وان الطرق والأساليب الأخرى من مقياس الميول ما زالت في طور الاستكشاف.

وقبل الخوض في العرض المفصل للقوائم المذكورة أعلاه لا بد من الإشارة إلى بعض النقاط التي يجب مراعاتها من قبل الأخصائي عند استخدام وتطبيق اختبارات الميول وهي:

1- إن تختار الاختبارات وفقاً لعمر وجنس الفرد الذي تطبق عليه.

2- يجب أن يختار الأخصائي الاختبار الذي يحقق ما يود أن يقيسه، فمثلاً يستحق استخدام اختبار كودر في البرنامج الجماعي كثير التفكير في الميول والاختيار المهني.

3- ينبغي أن تختار اختبارات الميول تبعاً لصدق وثبات المقاييس التي يتضمنها الاختبار والمعايير التي بني عليها، وهذا لأنه في بعض الأحيان يعتبر الاختبار صالحاً عندما يتم وضع معايير له على مجموعة محلية لتحقيق غرض خاص.

4- ينبغي أن تختار اختبارات الميول للحصول على نماذج ثابتة للميول المهنية.

5- يجب أن تختار اختبارات الميول التي يكون لها معايير يمكن الرجوع إليها ويجب أن تكون هذه المعايير صادقة.

قائمة سترونك للميول المهنية Strong Vocational Interest Blank SVIB

إن الشكل الحالي لهذه القائمة أنجز في عام 1966 ويحتوي على (399) فقرة متجمعة في ثمانية أقسام. يدون المفحوص في الأقسام الخمسة الأولى تفضيلا ته، وذلك بوضع دائرة حول واحد من الحروف (L الذي يشير إلى الرغبة) أو (I الذي

يشير إلى الاعتدال) أو (D يشير إلى عدم الرغبة). وكل واحد من الأقسام الخمسة يتعلق بواحد من الفئات التالية، المهن،المواضيع المدرسية، التسليات، النشاطات (مثل تصليح الساعة، إعداد خطاب، جمع الأموال الخيرية) وأنواع الأشخاص، أما الأقسام الثلاثة المتبقية من القائمة فتتطلب من المفحوص ترتيب النشاطات المعطاة حسب أفضليتها بالنسبة له، ومقارنة ميوله في أزواج من الفقرات، وتقدير قدراته الحالية وخصائصه الأخرى.

أن الهدف الأساس لهذا الاختبار-الذي يعتبر من أحسن الاختبارات المستعملة لقياس الميول-هو الكشف عن مدى اتفاق ميل فرد ما مع ميول الأفراد الآخرين الذين يشتغلون بمهنة معينة ممن بلغوا فيها درجة النجاح ويقوم الاختبار على أساس أن الأفراد في المهنة الواحدة تكون لهم نفس الميول، وهم بالتالي يختلفون في ميولهم عن الناس عموماً.

وكان الغرض الذي أقام عليه سترونك دراسته هو أن المجموعات المهنية-في ضوء ميولهم وحجم كراهيتهم،تفضيلهم أو عدم تفضيلهم-يمكن أن تتميز إحداها عن الأخرى. بمعنى أن أعضاء مجموعة مهنية ما (ولتكن الكيميائيين مثلاً) سوف تكون لهم مجموعة من الأشياء التي يحبونها أو يكرهونها،يفضلونها أو لا يفضلونها، والتي تختلف عن تلك التي يميل إليها أو لا يميل إليها أعضاء مجموعة مهنية أخرى مغايرة (كالمحامين مثلاً) ولإثبات هذا الغرض قام سترونك بمقارنة ميول مجموعات مهنية مختلفة، لا مباشرة إحداها مع الأخرى، وإنما بما اسماه (الأشخاص عامة) فمثلاً قارن ميول المحامين بميول الناس عامة وقارن ميول الكيميائيين بميول الناس عامة ...وهكذا، ووجد بأن ميول المحامي والكيميائي ليس فقط تختلف وإنما ذلك الاختلاف يكون بدرجات مختلفة كذلك، مما يكشف أيضاً عن اختلاف إحداهما عن الأخرى.

إن درجات القائمة تحسب بمقياس مختلف بالنسبة لكل مهنة، وحتى الآن تحتوي هذه القائمة على (54) مقياساً مهنياً بالنسبة لنموذج الرجال، و(32) مقياساً مهنياً

بالنسبة لنموذج النساء. أي أن لاختبار سترونك صورتان،صورة خاصة بالميول المهنية للرجال، وصورة خاصة بالميول المهنية للنساء. ومن وقت لأخر فإن عمليات تطوير تجرى على مقاييس أخرى جديدة، وبمرور الزمن يتم ضم مجاميع مهنية أخرى إضافية إلى المجاميع المهنية الأصلية للاختبار. وعند إضافة هذه المقاييس الجديدة، فإن الطريقة التي تضاف بموجبها هي نفس الطريقة التي أعدت بموجبها المقاييس المهنية الأولية للقائمة والتي تتلخص، بأن إجابات الأشخاص المشتغلين بنجاح في كل مهنة، تقارن بإجابات الرجال عامة (أو النساء عامة) أن المجموعة المرجعية الأساسية بالنسبة لنموذج الرجال، تضم رجالاً من تشكيلة واسعة من الأعمال التجارية والمهنية التي يدخلها خريجو الكليات عادة، وان اختيار المجموعة المرجعية يستند إلى حقيقة أن غالبية المقاييس المهنية لهذه القائمة تتعامل مع المهن والمواقف التجارية العليا. لذا فإن استخدام المجموعة المرجعية الممثلة لعينة الذكور الكلية أثبتت فعالية اقل في إظهار ميول متمايزة لمهنة الأفراد. فميول رجال التجارة والمهنيين كمجموعة تختلف كثيراً عن ميول العمال الماهرين، بينما نجد أن الاختلافات بين مهنة ذات مستوى عال وأخرى تكون غير واضحة عند استخدام المجموعة المرجعية الأكثر عمومية. إن المقاييس القليلة للمهن ذات المستوى المنخفض في هذه القائمة لا تكون متمايزة كما يجب إذا كانت المجموعة المرجعية مقتصرة على مهن ذات مستوى اقتصادي اجتماعي معين.

إن فقرات كل من المقاييس المهنية لهذه القائمة (SVIB) اختيرت وأعطيت لها الأوزان استناداً إلى الفروقات في تكرار الاختيار بين الرجال في تلك المجموعة المهنية والرجال عامة، ففي مقياس المحاماة على سبيل المثال، إن الوزن (+1) يشير إلى أن الإجابة على الفقرة كانت أكثر تكراراً بين المحامين عنه بين الرجال عامة، والوزن (−1) يشير إلى أن الفقرة كانت اقل تكراراً، أما الإجابات التي أخفقت في التمييز بين المحامين والرجال عامة فلم تظهر في مقياس المحاماة، بغض النظر عن مقدار تكرار اختيارها من قبل المحامين. إن المجموع الكلي للدرجة الخام الخاصة بالشخص

على كل من المقاييس المهنية هي ببساطة المجموع الجبري للأوزان الزائدة والناقصة التي حصل عليها بالنسبة لذلك المقياس، وهذه الدرجات الخام تحول إلى درجات معيارية بمتوسط حسابي قدره (50) وانحراف معياري قدره (10) ضمن حدود توزيع الدرجات في كل مجموعة مهنية كمحك، إن هذه المجموعات المحكية تضم عادة حوالي (300) شخصاً من الرجال الذين تتراوح أعمارهم بين (55,25) سنة، من الذين قضوا في المهنة المحددة ثلاث سنوات أو أكثر وكانوا مقتنعين بأعمالهم.

أن ما يحرزه الشخص من درجات في قائمة سترونج (SVIB) قد يكون لمهنة مستقلة. ولأغراض الاختبار، فأننا قد نرغب- على سبيل المثال- في معرفة كيف أن ميول طالب الوظيفة تشبه إلى حد بعيد ميول العاملين الناجحين في مكاتب التامين على الحياة. وغالباً ما يحرز المفحوص على درجات في جميع المقاييس المتيسرة في القائمة، وهذا يفيد في تحديد نمط الميل الكلي للفرد. وهذا النوع من التحليل يؤمن وسيلة ذات موثوقية كبيرة في التنبؤ بالميل المهني النهائي للفرد. وبالنسبة للإرشاد- الذي تجد فيه قوائم الميول تطبيقاتها الرئيسية - تعتبر دراسة الجوانب الكلية للميل مسألة جوهرية هامة. كما أن الإرشاد الإضافي المساعد يؤمن خلال استخدام المقاييس المهنية ذات النتائج المقاربة. وهذه المجموعات يمكن استنتاجها من خلال دراسة العلاقات بين المقاييس باستخدام التحليل العاملي والطرق الإحصائية الأخرى.

ومما تجدر الإشارة إليه في هذا الصدد أن (ثرستون) يعتبر أول من طبق التحليل العاملي على ثمانية عشر مقياساً مهنياً من مقاييس سترونج للميول المهنية وتوصل إلى تحديد أربعة عوامل هي:الميل العلمي، وعامل الميل اللغوي، وعامل الميل للأعمال التجارية،وعامل الميل إلى الناس كما وجد عدة عوامل خاصة أخرى.

وقد تتابعت البحوث الخاصة بالتحليل العاملي في ميدان الميول المهنية وكانت معظم هذه البحوث قائمة على اختبار سترونج. ولعل أوسع هذه البحوث هي سلسلة التحليلات العاملية التي قام بها سترونج نفسه والتي أدت إلى اكتشاف العوامل الآتية:

1- عامل الميل العلمي، يظهر في المهن العلمية-مثل الهندسة والطب والمهن المتعلقة بالطبيعة وعلم النفس.

2- عامل الميل المتعلق بالإعمال التجارية-وقد ظهر في مهن المحاسبة والتجارة.

3- عامل الميل إلى الأشياء في مقابل الأشخاص، ويظهر موجباً في المهن التي تتناول الأشياء بالتعديل والتغيير مثل التجارة والطباعة.

4- عامل الميل إلى اللغة ويظهر في المهن التي تتطلب استخدام اللغة مثل تدريس المواد الاجتماعية والمحاماة والتأليف.

5- عامل الميل للتعاون مع الناس ويظهر في المهن التي تهتم بالأشخاص مثل مديري المستخدمين ومديري المبيعات.

ويذكر بأن سترونك توصل إلى استنتاجه هذا من خلال القيام بإجراء أربعة دراسات عاملية على (25) (30) (22) (36) متغيراً على التوالي وقد كشفت هذه الدراسات عن اتفاق ظاهر الواحدة مع الأخرى ومع نتائج التحليل العاملي التي قدمها (ثرستون).

ويلاحظ أن المقاييس المهنية لقائمة سترونك جمعت في أحد عشرة مجموعة ضمت المقاييس المهنية الملائمة فيما بينها نسبياً من حيث الميول المعبرة عنهما، وعلى سبيل المثال فإن المجموعة الثانية (II) تشمل، المهندس المعماري، والمتخصص بالرياضيات والفيزيائي، والكيميائي، والمهندس. والمجموعة (الخامسةV) التي تتصف بوضوح الميل المشترك في النهوض أو الإصلاح الاجتماعي وتشمل مدير الملاك، مدير عام، مستشار إعادة التأهيل، أمين سر جمعية الشباب، باحث اجتماعي، معلم الاجتماعيات، ناظر مدرسة. وإضافة إلى ما ذكر فإنه توجد أربعة مقاييس إضافية أدخلت حديثاً، ولم توضع لحد الآن في مجموعات محددة.

وتتضمن قائمة سترونك إلى جانب المقاييس المهنية أربعة مقاييس غير مهنية

هي:مقياس المستوى التخصصي(SL) ومقياس المستوى المهني (OL) ومقياس الذكورة-الأنوثة (MF) ومقياس الإنجاز الدراسي (AACH). إن مقياس المستوى التخصصي اعد أصلاً بمقارنة إجابات الاختصاصيين في الطب بإجابات مجموعة الفيزيائيين عامة. وقد استحدثت مؤخراً ليكون قابل للتطبيق في ميادين أخرى للتعرف على الرجال الذين يمكن أن يواصلوا الدراسة المتقدمة من النمط الذي يتطلب تخصصاً دقيقاً.

وهناك من يشير إلى أن هذا المقياس يسمى بمقياس نضج الميل،وقد وضع هذا المقياس لتحديد ما إذا كان ميل المفحوص لمهنة ما ناضجاً نضج ميول الرجال الناجحين في هذه المهنة أو إنه غير ناضج مثله في ذلك مثل ميل الصغار في هذه المهنة. والدرجة المنخفضة على مستوى نضج الميل تعني أن ميول الفرد غير مستقرة نسبياً وان من الممكن أن تتغير بدرجة ملحوظة مع تقدم السن بينما الدرجة المرتفعة على هذا المقياس فتعني أن ميول الفرد قد وصلت إلى مستوى ثابت نسبياً، وانه إذا حدث تغير بعد ذلك فالاحتمال أن يكون تغيراً بسيطاً. ومن هنا يمكن أن تفيد من ناحيتين: الأولى معرفة مدى قرب ميل الشخص إلى النضج والثانية معرفة أي المهن يجب أن يتجه إليها أو يهتم بها وأيها لا يهتم، وذلك إضافة إلى ما نحصل عليه من درجات المقياس المهني.

ومقياس المستوى المهني (OL) يقيس الفرق بين ميول العمال غير الماهرين من جهة وميول الرجال المشتغلين بالمهن والتجارة من جهة أخرى. فإذا حصل شخص ما على درجة عالية على هذا المقياس كان معنى ذلك إنه يكشف عن بعض النشاط على المستوى الفني العالي أما إذا كانت درجة منخفضة كان معنى ذلك إنه يصلح للأعمال على مستوى اقل في المهارة.

ومقياس الذكورة - الأنوثة (MF) يبين هل أن ميول الفرد تشبه ميول الرجال والنساء. وقد وضع هذا المقياس على أساس الفرق في استجابات الرجال والنساء لفقرات اختبار الميول المهنية،وقد وجد سترونك أن من المفيد النظر إلى الدرجة على

هذا المقياس مع درجاته المهنية ليبين ما إذا كان الفرد يفضل القيام بمهنة تتصف بكثير من الميول الذكرية كالهندسة مثلاً أو بمهنة تتصف بكثير من الميول الأنثوية كالصحافة مثلا. فبعض المهن التي توجد لها درجات، يمكن أن تصنف بطريقة عامة باعتبارها مهناً ذكرية أكثر،أو مهناً أنثوية أكثر. فإذا كانت درجات الفرد عالية على الوظائف الذكرية وكان مقياس الذكورة-الأنوثة يشير عامة إلى الميول الذكرية،اتخذ ذلك بمثابة زيادة توكيد على صدق درجات الميول المهنية.

أما الإضافة الأكثر حداثة إلى المقاييس المهنية فهي مقياس الانجاز الدراسي (AACH) المطور من خلال مقارنة إجابات مجموعات من طلبة المدرسة العليا والجامعة ممن أحرزوا درجات دراسية عالية وممن احرزو درجات دراسية منخفضة. وعلى أساس البحوث التي أجريت بعد ذلك تبين أن هذا المقياس يعكس الميول في الأنشطة العلمية والعقلية مقابل الميول في الأنشطة العملية والتجارية التي تحتاج إلى مهارة. وقبل الانتقال إلى موضوع صدق وثبات قائمة سترونك لابد من الإشارة إلى أن الدكتور سيد محمد غنيم في كتابه سيكولوجية الشخصية ذكر بأن الدكتور عطية محمود هنا اعد (ترجم) الصورة الخاصة بالميول المهنية للرجال والاختبار يتكون من (400) فقرة.

بينما أشارت أنستا زي بأن فقرات الاختبار هي (399) مصنفة على النحو التالي: -

(100) فقرة تتصل بمقاييس المهن و(36) فقرة تتصل بالمواد الدراسية و(49) فقرة تتصل بأنواع التسلية و(48) فقرة تتصل بأنواع النشاط و(47) فقرة تتصل بأنواع الناس و(40) فقرة تتصل بالمفاضلة بين أنواع النشاط، و(40) فقرة تتصل بالمقارنة بين الميل إلى عملين و(40) فقرة تتصل بالقدرات الراهنة.

وفي الأقسام الخمسة الأولى، يطلب من المفحوص أن يضع علامة(صح) تحت الخانة الموضحة إلى جانب الفقرة والتي تشير إلى ما إذا كان يحب المهنة أو المادة أو لا

يحبه أو لا يهتم به. أما في القسم الخاص بالمفاضلة بين أوجه النشاط فيطلب من الشخص (في كل من المجموعات الأربعة التي يتكون منها هذا القسم والتي تحتوي كل مجموعة منها على (10) أنشطة أن يبين الأنشطة الثلاثة التي يفضلها أكثر من غيرها من أوجه النشاط العشرة. والثلاثة التي يفضلها اقل من غيرها ثم الأربعة الباقية التي تبقى محايدة.

وفي الجزء السابع الخاص بالمقارنة بين الميل إلى عمليتين مثلاً بين سائق سيارة عامة ومحصل في سيارة عامة. فإن الشخص في هذه الحالة عليه أن يبين هل يفض الأول أو لا يفضل احدهما على الآخر أم يفضل الثاني. وأخيراً في الجزء الثامن الخاص بتقدير القدرات والصفات الشخصية فعلى المفحوص أن يبين ما أذا كانت كل فقرة من الفقرات الأربعين التي يشتمل عليها هذا القسم يمكن أن تنطبق عليه أو لا تنطبق عليه أو إنه غير متأكد.

ثبات وصدق قائمة سترونك

لقد خضعت قائمة سترونك لبرنامج بحث متواصل، مما أدى إلى الحصول على معلومات واسعة عن ثباتها وصدقها. فمعاملات الثبات الفردية-الزوجية للمقاييس المختلفة أعطت درجة للثبات كانت بمعدل (0,88)وان واحداً فقط كان اقل من (0,80) وقد وجد بأن متوسط الارتباط بإعادة الاختبار على (139) من طلبة الصف الثاني في الكلية، والذين أعيد تطبيق الاختبار عليهم بعد فترة أسبوعين، وكذلك عند إعادته على (102) من البالغين بعد فترة ثلاثون يوماً كان (0,91). كما أظهرت الدراسات الطويلة الواسعة استقراراً جيداً طويل الأمد للدرجات،وهذه النتائج تضمنت الحصول على معامل ثبات قدره (0,68) عند إعادة تطبيق الاختبار بعد (3) سنوات بالنسبة لطلبة الصف الأول في الجامعة. ومعامل ارتباط قدره (0,61) عند إعادة الاختبار بعد (8) سنوات لطلبة الصف المنتهي في المدرسة الثانوية وعند إعادة تطبيق الاختبار بعد (22) سنة بالنسبة لطلبة السنة الأخيرة في

الكلية بلغ معامل الارتباط ((0,67)) ويلغ معامل الارتباط (0,56) عند إعادة الاختبار على (48) من مدراء البنوك بعد (30) سنة وكان التطبيق الأول عندما كانت أعمارهم (40) سنة تقريباً. وعموماً،فإن درجات الميول تكون اقل استقراراً بالنسبة للطلبة في مستوى المدرسة الثانوية، إلا أن مثل هذه الدرجات تكون ذات استقرار جدير بالاعتبار بالنسبة لطلبة الجامعة.

وهناك نمط آخر من التحليل يتعلق باستقرار متوسط المظاهر المهنية بمرور الزمن. وبالنسبة لهذا الغرض،فإن قائمة سترونك طبقت على عينات مختلفة من الرجال المهيمنين على نفس الأعمال في نفس التنظيمات المهنية في فترتين منفصلتين بلغت المدة الزمنية بين التطبيق الأول وإعادة تطبيقه (30) سنة أو أكثر. وهذه الطريقة اتبعت مع أربع مجموعات مهنية: نظراء المدارس، مد راء البنوك، ورؤساء النقابات. وبشكل عام فإن متوسطات المظاهر المهنية عبر هذه الفترة الزمنية كانت متشابهة بشكل ملفت للنظر. وذلك يوحي بأن المقاييس التي أعدت على المجموعات المحكية الصالية مازالت قابلة للتطبيق في الوقت الحاضر.

وفيما يتعلق بالصدق فهناك دليل يسترعي الاهتمام بأن الأشخاص يميلون إلى الالتحاق والبقاء في المهن التي يحرزون في المقاييس العبرة عنها في قائمة سترونك درجات عالية. وواحدة من أكثر الدراسات الطولية اتساعاً هي تلك التي قام بها سترونك نفسه، إذ تابع بعد (18) سنة (663) من طلبة جامعة ستانفورد من الذين طبقت عليهم القائمة عندما كانوا في الكلية. وقد وجد بأن هناك ((88)) فرصة من أصل (100) بأن الطلبة الذين أحرزوا (55) درجة أو أكثر على احد المقاييس المهنية سوف يعملون في تلك المهنة. بينما هناك (17) فرصة من أصل (100) بأن الطلبة الذين أحرزوا اقل من (30)درجة سوف يعملون في تلك المهنة.

وبالنسبة لطلبة الثانوية فإنه يمكن التنبؤ على نحو جيد بالحقل أو المجموعة المهنية المعبرة عن الوظيفة النهائية التي سيشغلونها، شريطة أن تكون إجاباتهم على القائمة تعبر عن أنماط ميولهم المتميزة، والمعبرة عن إجابات حقيقية. وهناك أيضاً

بيانات توضح العلاقة بين الدرجات على قائمة سترونك، والنجاح في العمل المناسب أو الرضا عن العمل، إلا إن الدليل بالنسبة للصدق التنبؤي مقابل هذه المحاكات هو ضئيل وتفسيره أكثر صعوبة،وعموماً فإن الاستمرار في مهنة معينة يؤمن دليلاً جيد على أن الفرد أحرز على الأقل أدنى مستوى للنجاح وانه وجد العمل المناسب له إلى حد معقول.

أن مشاريع بحوث متعددة جارية حالياً على هذه القائمة في المركز الخاص بقياس الميل الذي أسس سنة (1962) في جامعة منيسوتا. وأحد هذه المشاريع هو الدراسة الطولية لفترة (35) سنة على(1000)من الطلبة بعمر (15) سنة وكان التطبيق الأول عام (1930) وكل شخص في هذه الدراسة أعيد تطبيق القائمة عليه مرة ثانية بعمر (50) سنة تقريباً مع تقديم خلاصة لتاريخه المهني. وهكذا فإن هذه الدراسة قدمت بيانات مختارة تتعلق بكل من الاستقرار طويل الأمد للدرجات والصدق التنبؤي للقائمة.

ومما تجدر الإشارة إليه بأن قائمة ميل مشابهة أعدت بالنسبة للمهن المتنوعة التي تتطلب مهارة تامة أو متوسطة، وذلك بعد قائمة سترونك، إلا أنها استخدمت التجار عموماً كمجموعة مرجعية. وهذه القائمة هي قائمة منيسوتا للميل المهني التي وفرت (21)مقياساً لمهن مثل خباز، أو سائق سيارة الحليب، سمكري، ومصلح الراديو والتلفزيون. وهناك أيضاً تسعة مقاييس تظهر رغبة المفحوص للنشاطات العملية الممثلة بالمهن المختلفة كالميكانيكي،والخدمة في المطاعم،والالكترونيات. إن الفقرات في قائمة منيسوتا للميل المهني بثقل اكبر نحو النشاطات المهنية عنه في قائمة سترونك. وأنها عرضت وفقاً لأسلوب الاختيار الإجباري من النوع الثلاثي. كما أن التمايز بين المهن في هذه القائمة ليس قاطعاً (حادا) كما في سترونك. ومن المحتمل أن يكون الاختلاف بين هاتين الوسيلتين ناجحاً عن طبيعة المهن نفسها.

بيان المفاضلة المهنية لكودر

وهذا الاختبار وضع بعد اختبار سترونك، وقد اتبع فيه كودر أسلوباً مختلفاً في اختيار وتقديرات الفقرات. وإذا كان اختبار الميول المهنية لسترونك يمثل اتجاهاً تجريبياً أو اختبارياً في بناء المقاييس، فإن اختبار التفضيل المهني لكودر يمثل اتجاهاً عقلياً يقوم على أسس نظرية محددة ويتجه نحو أهداف معينة دون الاستناد إلى الخبرة أو الحقائق المشتقة منها. والهدف الأساس للاختبار هو الكشف عن الميل النسبي في عدد قليل من المجالات الواسعة أكثر مما هو في الكشف عن مهن محددة.

وقد لاحظ كودر أن اختبار سترونك، هو من الاختبارات التي تحتاج إلى وقت طويل وجهد كبير وان وضع مقاييس مهنية تغطي المهن المختلفة الموجودة حالياً أمر يكاد يكون مستحيلاً. فقد استغرق سترونج مايزيد على (25)سنة في وضع (60)مقياساً تغطي جانباً محدوداً من آلاف المهن الموجودة. كما لاحظ أيضاً إن بعض المقاييس الموضوعة لقياس مهن معينة يرتبط إحداها بالآخر ارتباطاً عالياً. ومعنى ذلك أن النتائج التي نحصل عليها من معرفتنا بدرجات الفرد على احد المقاييس لن تزيد كثيراً بمعرفتنا بدرجاته على المقياس الآخر، فليس ثمة كسب كبير يمكن أن يعوض الجهد والوقت الذي يبذل في وضع المقياس، والقيام بتقديره بعد ذلك. ومن هنا كان احد أهداف كودر أن يؤكد منذ البداية أن أي مقياس جديد يجب إلا يرتبط بمقياس آخر سبق وضعه ارتباطاً عالياً. حتى لا نكرر النتائج التي سبق الوصول إليها من قبل. فهو يريد مقاييس لا يرتبط أحدها بالآخر. وفقرات الاختبار صيغت أصلاً وجمعت (تجريبياً) على أساس صدق المحتوى،وقد اتبع هذا من خلال التحليل الشامل للفقرات على طلبة المدرسة الثانوية ومجموعات الراشدين والهدف من وراء هذا التحليل هو لإظهار مجموعات الفقرات ذات الاتساق الداخلي العالي وذات الارتباط المنخفض مع المجموعات الأخرى للفقرات. وقد تحقق هذا الهدف بشكل جيد ومعقول بالنسبة لأغلب مقاييس الاختبار.

إن اختبار كودر هذا وضع لقياس ميول التلاميذ من سن (9 سنوات إلى 19 سنة) ويصف ميول التلاميذ في ميول مختلفة، ويفيد في التوجيه المهني.

خطوات بناء الاختبار

اتبع كودر الخطوات الآتية في بناء اختباره:

1- إعداد قائمة تتكون من (200) وجهاً من أوجه النشاط. وهذه الأنشطة بدت - على أساس قبلي- مفيدة كأدلة على تفضيل الميل. وقد رتبها كودر في (40) مجموعة كل منها يتألف من خمسة أنشطة مختلفة بحيث تكون أوجه النشاط ممثلة في كل مجموعة وأعطى هذه الصورة من الاختبار إلى (500) طالب بالجامعة وطلب إلى كل منهم أن يرتبها حسب تفضيله إياها. وذلك في كل مجموعة من المجموعات الأربعين.

وبذا بدت مجموعة من أوجه النشاط هذه قابلة لان تصنف باعتبارها ميكانيكية في طبيعتها. بينما بدت مجموعة أخرى أنها قابلة لان تصنف باعتبارها أدبية في طبيعتها. وباستخدام هذه الفقرات- على أساس قبلي- حدد أوزان هذه الفقرات ليبين الميل للعمل الميكانيكي وتفضيله على الأنواع الأخرى من النشاط. والميل للعمل الأدبي وتفضيلها على الأنواع الأخرى من النشاط. وقد وجد كودر أن الثبات بالنسبة لفقرات المقياس الأدبي باستخدام طريقة التجزئة النصفية هو (0.85) ومن ثم اعتبره مقياساً ثابتاً بدرجة معقولة لشيء واحد هو تفضيل النشاط الأدبي. واتخذ كودر هذا المقياس نقطة ارتكاز لوضع غيره من المقاييس.

2- حسب كودر معاملات الارتباط بين استجابات كل فقرة من فقرات الاختبار المائتين والدرجة الكلية لمقياس الميل الأدبي. وعزل الفقرات التي وجد أنها ترتبط ارتباطاً منخفضاً بالميل الأدبي، وقام بدراستها. وكشفت دراسة محتوى هذه الفقرات عن وجود عدد كبير من الفقرات الدالة على تفضيل النشاط

التجريبي أو العملي والذي حدده كودر بعد ذلك باسم (الميل للعمل العلمي) وقد كون (كودر) من هذه الفقرات أساس مقياس العمل العلمي، وحسب معامل ثباته بطريقة التجزئة النصفية ووجد إنه حوالي (0,65).

3- فحص كودر الفقرات غير المتضمنة في المقياس الأدبي والعلمي. واختار منها تلك التي تبدو أنها تشير إلى تفضيل النشاط الفني. وقد حدد الارتباط بين أوجه النشاط المهنية وبين هذا المقياس الجديد. وأضاف إلى المقياس أوجه النشاط التي ارتبطت به ارتباطاً عالياً. وعند إضافة هذه الفقرات حاول كورد أن يوازن قدر الإمكان بين ارتباط هذه الفقرات والمقياس الأدبي والعلمي. وذلك بجعل معاملات الارتباط بينهما قريبة من الصفر وعلى نحو ما فعل بالنسبة للمقياسين السابقين حيث ظلت الارتباطات بينهما أيضاً قريبة من الصفر.

4- وكان المقياس الرابع الذي وضعه كودر يهدف إلى قياس المركز الاجتماعي. واتبع فيه نفس الأسلوب الذي اتبعه في المقاييس السابقة ولكن المشكلة هنا كانت أكثر تعقيداً. فالفقرات يجب أن لا يكون بينهما أي ارتباط والمقاييس الأدبية والعلمية والفنية. وبعد تكملة مقياس المركز الاجتماعي ظهر أن الفقرات المتبقية لا يمكن تصنيفها في مقاييس أخرى ولذلك أضاف أوجه نشاط جديدة، ولكنه وجد أنها ترتبط إلى حد ما مع هذا المقياس الأخير. ولذلك اسقط كودر مقياس المركز الاجتماعي ووزع كثير من فقراته على مقياسين آخرين اقترح وجودهما. وتابع نفس الأسلوب من العمل إلى أن وصل إلى بناء سبعة مقاييس هي، الميل للعمل الأدبي والعلمي والفني والحسابي والإقناعي والموسيقي والخدمة الاجتماعية.

وقد نشر كودر الاختبار بصورته هذه،ولكن نتيجة للنقد والدراسة شعر كودر بضرورة إضافة مقاييس الميل للعمل الميكانيكي والكتابي (وقد وضع كودر هذين المقياسين في ضوء معيار ثباتهما الداخلي دون أن يهتم ببيان كيف ترتبط فقراتهما

بالدرجات الكلية على المقاييس السبعة الأخرى) وأخيراً أضيف إلى الاختبار مقياس العمل الخلوي (أو العمل في الخلاء).

ويمكن القول بوجه عام أن كودر قد نجح إلى حد معقول في الوصول إلى هدفه وهو وضع مقاييس ثابتة من متغيرات مستقلة تقريباً. وإذا كانت هذه المقاييس تغطي الميول بطريقة مناسبة. فإنها سوف تكون ذات قيمة كبيرة من الناحية الاقتصادية إعطاء الشخص فكرة عن مجالات اهتمامه واتجاهات ميوله.

وقد أورد كودر قوائم مفتوحة بالمهن التي تعتبر مناسبة لهؤلاء الذين يحصلون على درجات عالية على كل مقياس من المقاييس العشرة. وقد وضعت هذه القوائم أولاً على أساس منطقي، ولكن بعد تراكم المادة التجريبية،تم إدخال بعض التعديلات عليها.

ويتألف اختبار كودر هذا من (160) فقرة تتصل ببعض أساليب النشاط، والفقرات مصممة بطريقة الاختيار الإجباري من النوع الثلاثي إذ يعبر عن كل فقرة بثلاثة أنشطة ملائمة،وما على المفحوص إلا أن يشير إلى أي من هذه النشاطات يفضله أكثر ما يمكن واي منها يفضله اقل ما يمكن وبالطبع فإن النشاط المتبقي سوف يحتل بالنتيجةالمرتبة الوسط بين النشاطين المختارين. وعند التقدير يعطى النشاط المفضل (درجتان) والنشاط غير المفضل (صفر) أما النشاط المتبقي فيعطى (درجة واحدة) لأنه يقع في الوسط والملاحظ في الفقرات أن كل منها يعبر عن أكثر من مجال واحد من المجالات المهنية ويصل في أحيان كثيرة إلى ثلاثة مجالات وهو الحد الأقصى الذي يمكن أن تعبر عنه الفقرة الواحدة بأنشطتها الثلاثة.

وقبل القيام بعملية تصحيح الاختبار، لابد من الإشارة إلى أن كودر وضع إلى جانب المقاييس العشرة المذكورة سابقاً لقياس الميول مقياساً إضافيا هو "مقياس التحقق" وهدفه اكتشاف الإجابات غير المدروسة والناتجة عن سوء الفهم أو الاختيار المبني على أساس الرغبة الاجتماعية. وذلك لاستبعاد الإجابات المعبرة

عن الميول الزائفة حيث ظهر أن التزييف بحدود معينة في اختباري سترونك وكودر. إلا أن إمكانيته اكبر في اختبار كودر، وذلك يعود إلى الطبيعة الأكثر وضوحاً لفقراته.

وبعد التحقيق من صدق الإجابة باستخدام المقياس المذكور، تستخرج الدرجات الخام الخاصة بكل ميل من الميول، وفقاً لتعليمات خاصة بذلك عن طريق مفاتيح خاصة بالاختبار، ثم تترجم كل درجة خام إلى المقابل المئوي لها ثم يرسم بروفيل أو التخطيط العام للفرد في مختلف الميول. والتفسير المهني يتم عادة بالتعرف على أعلى تقديرين في البروفيل. واللتين تشيران إلى قائمة من المهن التي يعتقد بأن هذه التقديرات تناسبها، وبالإضافة إلى ذلك فإن المجالات التي تشير إلى أدنى اهتمام أو ميل لدى الفرد تعتبر أيضاً ذات أهمية إذ تعني أن الفرد لا يحب العمل الذي يتطلب مثل هذا النشاط.

وبالإضافة إلى أن كودر وضع تحت كل مجال العديد من المهن الخاصة التي ترتبط بالميل الرئيسي لذلك المجال، والتي يجب أن تلقى اهتماماً من الفرد،فإنه ذكر أيضاً عدد من المهن التي ضمن زوج من مجالات الميول مثال ذلك المهن التي تعبر عن الثنائيات التالية للميول،ميكانيكي- فني، ميكانيكي- علمي، علمي- فني، علمي- خدمة اجتماعية، إقناعي- أدبي، خلوي- فني. فمثلاً. مصلح الراديو يصنف تحت مجال الميول الميكانيكية،بينما مهندس المناظر تحت مجال الميول الخلوية- الفنية. وان كل هذه الأمور موضحة في كتيب التعليمات الخاص بالاختبار.

ثبات وصدق اختبار كودر

إن معاملات الثبات لمقاييس كودر، كما حددت باستخدام طريقة كودر- ريجاردسون مقاربة ل(0,90) وان معاملات الاستقرار(الثبات) لفترة سنة أو اقل كانت عالية أيضاً. إلا أن المعلومات المتيسرة عن معاملات الثبات عبر فترات زمنية أطول قليلة جداً. وهنالك بعض الأدلة التي تشير بأن تغيرات نسبية تحدث في

مجالات الميول العالية والمنخفضة - وخاصة بالنسبة لطلبة المدرسة الثانوية-عندما تجرى إعادة الاختبار بعد عدد من السنوات المتباعدة.

أما بحث الصدق بالنسبة لهذه القائمة، فغالباً ما يتم التوصل إليه بالمقابلة مع محك الرضا عن العمل. وفي واحدة من أكثر الدراسات الطولية شمولية. تم توجه استبيان الرضا عن العمل إلى (1164) شخصاً ممن طبق عليهم بيان المفاضلة المهنية لكودر خلال دراستهم في المدرسة الثانوية،وذلك بعد عشر سنوات من تطبيق الاختبار الأول عليهم- أي اختبار كودر- وفي هذا الوقت ظهر بأن(728) شخصاً منهم كانوا مشتغلين بأعمال صنفت على أنها متسقة مع أنماط ميولهم الأصلية، و(436) كانوا مشتغلين بأعمال صنفت على أنها غير متسقة مع تلك الأنماط. وقد كانت النسبة المئوية للمقتنعين بأعمالهم في المجموعة المتسقة (62٪) بينما نسبتهم في المجموعة غير المتسقة كانت (34٪) ومن جهة أخرى فإن النسبة المئوية للعمال غير المقتنعين كانت (8٪) للمجموعة المنسقة و(25٪) للمجموعة غير المتسقة.

إن نتائج مشابهة تم التوصل إليها في دراسات متعددة للأشخاص المستخدمين في ميادين محددة، كرجال الدين، والعمال الصناعيين، ومرشدي إعادة التأهيل المهني، والمحاسبين. وعلى سبيل المثال، فإن مجموعة المقتنعين بين المحاسبين حصلوا على درجات أعلى من مجموعة غير المقتنعين في مقاييس الميول الحسابية والكتابية وبشكل ذو دلالة، إلا أن درجاتهم على مقاييس الميول الفنية والعلمية والخلوية كانت منخفضة وبشكل ذو دلالة.

مسح الميل العام لكودر Kuder General Interest Survey

لقد طور هذا الاختبار حديثاً جداً كتنقيح وتوسيع لبيان المفاضلة المهنية لكودر- سابق الذكر- وهو مصمم للمراحل الدراسية من (6-12) مستخدماً لغة مبسطة ومفردات سهلة مقارنة بقائمة كودر الأولى، حيث يكفي المستوى القرائي للمرحلة الدراسية لفهم محتويات القائمة الجديدة المطورة. لقد تم التوصل إلى

المعايير المئوية لهذه القائمة من خلال استخدام عينة طبقية شاملة تتضمن أكثر من عشرة آلاف ولداً وبنتاً وعينات صغيرة من الراشدين والبحث مازال جار لاكتشاف ما إذا كانت هذه القائمة ملائمة للراشدين على حد سواء كما هي بالنسبة للأصغر سناً وفي أي الحالات يمكن أن تحل هذه القائمة نهائياً محل بيان المفاضلة المهنية الأول لكودر وعموماً فإن معاملات الارتباط بين درجات القائمتين عالية بقدر ارتفاع معاملات ثباتهما تقريباً. لقد تم بناء القائمة الجديدة من خلال استخدام criterion Keying Procedure مشابهة لتلك المتبعة في قائمة سترونك إلا أنها لم تستخدم المجموعة المرجعية للرجال أو للنساء عموماً وبدلاً عن ذلك اعتبرت درجة المفحوص على أي من المقاييس المهنية كتعبير للارتباط بين نمط ميله ونمط الميل السائد في المجموعة المهنية المثبتة. إن حساب درجات هذه القائمة (تصحيحها) لا يمكن إجراءه يدوياً. إذ تعاد صفحات الجواب إلى مطبق الاختبار (ناشره) ويتم التصحيح من خلال الربط بين عمل جهاز الفرز البصري السريع والحاسبات الإلكترونية.

إن الدرجات في هذه القائمة تحسب بالنسبة لـ(79) مهنة و(20) حقلاً كلياً رئيساً (College major fields) بالنسبة للرجال و(56) مهنة و(25) حقلاً كلياً رئيساً بالنسبة للنساء. وعلى أية حال فإنه أكثر من ثلث درجات النساء تحرز على مقاييس معدة لمجموعات محكية خاصة بالذكور في ميادين، الرجال هم المهيمنين فيها إلا أنها مع ذلك تقدم فرصاً مهنية مناسبة للنساء. إن المهن التي تغطيها هذه القائمة متباينة على نحو واسع في المستوى. فهي تمتد مَن الخباز وسائق الشاحنة إلى الكيميائي والمحامي والأمر الذي أدى إلى هذه التغطية الواسعة للمهن وبوسيلة واحدة فقط هو إهمال أو عدم استخدام المجموعة المرجعية. وبالإضافة إلى المقاييس المهنية والمقاييس الكلية الرئيسية فإن هذه القائمة تؤمن مقياساً آخر وهو مقياس التحقق وثمانية مقاييس تجريبية مستخدمة حالياً لإغراض البحث فقط وبينما نجد أن هذه القائمة مصممة للتطبيق على الأفراد من المرحلة الدراسية العاشرة إلى مستوى الراشدين، إلا أنها تتطلب مفردات قرائية بمستوى المرحلة السادسة تقريباً وكما في قوائم كودر

الأخرى فإن فقرات هذه القائمة أعدت وفقاً لأسلوب الاختيار الإجباري من النوع الثلاثي ومحتوى الفقرات يغطي مجالات واسعة من الميول متضمنة أصلاً في القائمة الأولى سالفة الذكر، إضافة إلى مجالات أخرى متعددة، كالمفاضلة في اختيار العمل الاستقلالي وغيرها.

ومن خلال التحليل الإحصائي المركز لدرجات (3000) شخص و(100) شخصاً في كل مجموعة من المجموعات الثلاثين الممثلة بشكل جوهري للمهن والميادين الكلية الرئيسية التي تغطيها القائمة اظهر كودر بوضوح بأن أفضل تمايز بين المجموعات المهنية يمكن الحصول عليه باستخدام نظام إعطاء الدرجات الخاص بهذه القائمة، عنه في حالة استخدام المقاييس المهنية المشتقة من استخدام المجموعة المرجعية للرجال عموماً. كما أن المقاييس الثلاثون استخدمت نفسها في التحليل الواسع للثبات بإعادة الاختبار، وللارتباط المتداخل للدرجات على المقاييس المختلفة، والجوانب الفنية الأخرى للقائمة، وفي كل هذه التحليلات فإن هذه القائمة ظهرت لتكون مرضية جداً.

أما البحث المتعلق بالمقاييس المهنية الأخرى، إضافة إلى بحث الصدق مقابل محكات خارجية فإنه ما زال ينشد الانجاز. وعموماً فإن هذه الوسيلة أظهرت بعض التجديدات الأساسية في بناء الاختبار، وأساسها الحقيقي والجوهري يمكن أن تحدد قيمته وذلك بعد القيام بالبحث المضاعف والتطبيق في واقع السياقات العملية.

اختبار مسح الميول ليلفورد وزيمومن

وهو اختبار يغطي قطاعاً كبيراً من الميول في عدد قليل من المجالات ويتكون الاختبار من نشاطات كثيرة ومتنوعة ويحتوي الاختبار على تسعة مجالات واسعة وكل قسم منها ينقسم إلى فرعين وكالآتي:

فني: تقديري، تعبيري

لغوي: تقديري، تعبيري

علمي: بحث علمي، نظري

ميكانيكي: تنفيذ، تعميم

خلوي: طبيعة، رياضة بدنية

الأعمال السياسية: تجاري، زعامة

النشاطات الاجتماعية: إقناعي، اجتماعي

المساعدة الفردية: الخدمات الفردية، الخدمات الاجتماعية

العمل الوظيفي: الكتابي، العددي

وقد قسمت المجالات على أساس الارتباط والتداخل فيما بينها داخل المجال الواحد ويميز المقياس بين النشاطات الترويحية والنشاطات التي يتخذها الأفراد كحرفة،أي أن المفحوص يؤشر أمام كل فقرة فيما إذا كان يحب هذا النشاط كحرفة أو كهواية أو كليهما معاً، أو لا هذه ولا تلك، لذلك يوضع درجتان، درجة للمقياس المهني والأخرى للمقياس الترويحي.

1- الثبات:

وقد استعمل الباحثان طريقة التجزئة النصفية لاستخراج معامل الثبات،وكان معامل الثبات عالياً يتراوح من 0.80 – 0.94

2- الصدق:

يتمتع هذا الاختبار بنوع من الصدق المنطقي وهو صدق المحتوى Content Validity الذي اتبعه الباحثان لتقرير صدق هذا النوع من الدراسة.

اختبار الميول المهنية لأحمد زكي صالح

يتضمن مجموعة من النشاطات المهنية بحيث يمكن استخدام هذا الاختبار في التوجيه المهني، وهو نفس اختبار كودر إلا إنه ترجم إلى اللغة العربية وقنن على البيئة المصرية.

ويتكون الاختبار من (504) فقرة صنفت إلى مجاميع وكل مجموعة تحتوي على ثلاثة فقرات من مجالات الميول المختلفة وعلى المفحوص أن يختار أكثر أساليب النشاط تفضيلاً واقلها تفضيلاً، أما الميول فهي:

1- الميل الخلوي
2- الميل للعمل الميكانيكي
3- الميل الحسابي
4- الميل الإقناعي
5- الميل العملي
6- الميل للعمل الفني
7- الميل للعمل الأدبي
8- الميل الموسيقي
9- الميل للخدمة الاجتماعية
10- الميل الكتابي

كفاءة الاختبار

1- **ثبات اختبار الميول:** اتبع الباحث طريقة التجزئة النصفية لإيجاد الثبات باستخدام معادلة كودر ريتشارد سون، وتراوح معامل الثبات ما بين 0.62–0.99.

2- **الصدق:** طبق هذا الاختبار على مجموعات من خريجي الكليات واستخرجت المتوسطات لهذه المجموعات المهنية، وترجمت هذه المتوسطات إلى المعايير المئوية لها. وقد كان هناك دلالة معنوية عند مستوى 1٪ مما يدل كل صدق المقياس.

قائمة المصادر

المصادر العربية

الأمام، مصطفى محمود وآخرون، التقويم والقياس، دار الحكمة، بغداد، الطبعة الأولى.

أبو حطب، فؤاد، وسيد احمد عثمان، التقويم النفسي، القاهرة، مكتبة الأنجلو المصرية.

أبو حطب، القدرات العقلية، دار الكتب الجامعية، بيروت، لبنان، 1987.

أبو لبدة، سبع، مبادئ القياس النفسي والتقويم التربوي، الأردن، عمان.

احمد، محمد عبد السلام، القياس النفسي والتربوي، مكتبة النهضة المصرية، القاهرة.

البطش وآخرون، دليل الباحث في المقاييس النفسية والتربوية في ميدان القدرات والشخصية، الجامعة الأردنية، مكتبة العلوم التربوية، برنامج البحث التربوي، والخدمات التربوية والنفسية، الجزء الأول 1994.

اسعد، ميخائيل، الإحصاء النفسي وقياس القدرات العقلية، الطبعة الأولى، دار الآفاق الجديدة، بيروت.

اميول، أي، أي، ترجمة صبحي عبد اللطيف المعروف وآخرون، الأسس النفسية في التربية، عالم المعرفة.

أنستا زي، ترجمة السيد محمد خيري، سيكولوجية الفروق بين الأفراد والجماعات، الجزء الأول، القاهرة، الشركة العربية للطباعة والنشر، 1959.

بركات، محمد خليفة، تحليل الشخصية، الطبعة الثالثة، القاهرة، مكتبة مصر.

البسام، عبد العزيز، دليل السجل المجمع لشخصية التلميذ، وزارة التربية.

التميمي، خالد بن حسن شيبان، اثر كل من نوع المحكم وطول الاختبار على تحديد درجة القطع لاختبار محكي المرجع يقيس الكفايات الرياضية في العمليات الحسابية على الأعداد بالصف السادس الابتدائي بمدينة جدة، جامعة أم القرى، كلية التربية، المملكة العربية السعودية، رسالة ماجستير غير منشورة، 1999.

توق، عدس، علم النفس التربوي، دار الشروق، 1984

تيسير وآخرون، أسس الإدارة التربوية والمدرسية والإشراف التربوي، عمان، دار للنشر والتوزيع.

ثورندايك، روبرت وآخرون، ترجمة عبد العزيز عبد الرحمن عدس، القياس والتقويم في علم النفس، مركز الكتاب الأردني، عمان.

جابر عبد الحميد جابر وآخرون، مناهج البحث في التربية وعلم النفس، الطبعة الثانية، دار النهضة العربية، القاهرة.

جابر، عبد الحميد جابر، مدخل لدراسة السلوك الإنساني، دار النهضة العربية، الطبعة الأولى.

جابر، عبد الحميد جابر، الذكاء ومقاييسه، القاهرة دار النهضة العربية، الطبعة الرابعة، 1977

الجسماني، عبد علي، علم النفس وتطبيقاته التربوية والاجتماعية، مكتب الفكر العربي، بغداد، الطبعة الأولى.

جلال، سعد، القياس النفسي، مكتبة المعارف الحديثة، الإسكندرية، الطبعة الأولى.

جمال، مفيد حسن وآخرون، تمايز العمر والجنس في الميول الاجتماعية، المجلة المصرية للدراسات النفسية، العدد السادس، القاهرة.

الحارثي، زايد عجير، بناء الاستفتاءات وقياس الاتجاهات، جدة، دار الفنون، 1992.

حبيب، مجدي عبد الكريم، التقويم والقياس في التربية وعلم النفس، مكتبة دار النهضة المصرية، 1996.

خير الله، سيد محمد، وآخرون، القدرات ومقاييسها، القاهرة، مكتبة إلا نجلو المصرية.

خيري، السيد محمد، الإحصاء في البحوث النفسية والتربوية والاجتماعية، الطبعة الرابعة، دار النهضة العربية، القاهرة.

دافيدون، لندال ترجمة سيد الطواب، مدخل علم النفس، دار ماكروهيل للنشر الطبعة الثالثة، 1988.

داود، عزيز حنا وآخرون، دراسات في علم النفس، الجزء الأول، القاهرة، مكتبة النهضة المصرية.

داود، عزيز حنا وآخرون، مناهج البحث التربوي، وزارة التعليم العالي، بغداد، الطبعة الأولى.

داود، عزيز حنا وآخرون، الشخصية، وزارة التعليم العالي، جامعة بغداد، الطبعة الأولى.

الدوسري، ابراهيم مبارك، إطار مرجعي في القياس والتقويم التربوي، المركز العربي للبحوث التربوية لدول الخليج العربي، الرياض، 1999.

الرواشدة وآخرون، مرشد المعلم في بناء الاختبارات التحصيلية، وزارة التربية، المديرية العامة للامتحانات، عمان، 2000

الروسان، فاروق، أساليب القياس والتشخيص في التربية الخاصة، دار الفكر للطباعة والنشر والتوزيع، الطبعة الأولى.

الرفاعي، طاهرة عيسى خلف، بناء مقياس للميل العلمي والأدبي، رسالة ماجستير، كلية التربية، ابن رشد 1978.

زهران، حامد عبد السلام، علم النفس الاجتماعي، الطبعة الرابعة، عالم الكتب، القاهرة.

زهران، حامد عبد السلام وآخرون، استفتاء القيم، جامعة عين شمس، القاهرة.

الزوبعي، عبد الجليل وآخرون، الاختبارات والمقاييس النفسية، العراق، جامعة الموصل، 1980

الزوبعي، عبد الجليل وآخرون، علم النفس التربوي، معاهد المعلمين، مطبعة وزارة التربية.

زيدان، محمد مصطفى وآخرون، علم النفس، القاهرة، دار القرآن للطباعة

سالم، مهدي محمود، الأهداف السلوكية، الطبعة الأولى، مكتبة العبيكان، الرياض 1997.

سمارة، عزيز، وآخرون، مبادئ القياس والتقويم في التربية دار الفكر للنشر والتوزيع، عمان، الأردن، 1989.

سويف، مصطفى، مقدمة لعلم النفس الاجتماعي، القاهرة

السيد، يسري مصطفى، مبادئ ومفاهيم أساسية في القياس والتقويم، جامعة الإمارات العربية المتحدة، كلية التربية، مركز الانتساب الموجه، أبو ظبي، 2004.

السيد، يسري مصطفى، ورشة عمل حول صياغة الأهداف التعليمية، جامعة الإمارات العربية المتحدة، كلية التربية، مركز الانتساب الموجه، أبوظبي، 2004.

السيد، محمد توفيق، بحوث في علم النفس، القاهرة، المطبعة الفنية الحديثة

السيد، فؤاد البهي، علم النفس الاجتماعي، رؤية معاصرة، القاهرة، دار الفكر العربي

السيد، فؤاد البهي، الذكاء، الطبعة الرابعة، القاهرة، دار الفكر العربي، 1976

السيد، فؤاد البهي، علم النفس الإحصائي وقياس العقل البشري، الطبعة الثانية، دار الفكر العربي، القاهرة.

الطائي، نزار مهدي مقارنة بين مستويات نمو الميول المهنية للشباب في الجمهورية العراقية وجمهورية مصر العربية، رسالة ماجستير غير منشورة.

الطريري، عبد الرحمن بن سليمان، القياس النفسي والتربوي، الرياض، مكتبة الرشد 1997.

العاني، نزار محمد سعيد، اختبارات الذكاء، محاضرات مطبوعة بالرونيو، 1984.

العمر، علاء كامل صالح، اثر الزيارات الميدانية للمؤسسات الصناعية في تنمية الميول المهنية لطلاب المرحلة المتوسطة، رسالة الدكتوراه غير منشورة.

عبد الرحمن، سعد، القياس النفسي، الكويت، مكتبة الفلاح، 1997.

عبد ارحمن، سعد، القياس النفسي، النظرية والتطبيق، دار الفكر العربي، القاهرة، 1998.

عبد الخالق، احمد، قياس الشخصية، دار القلم، الكويت.

عبد الغفار، عبد السلام وآخرون، علم النفس الاجتماعي، القاهرة، دار النهضة العربية.

عبد المعطي، عبد الباسط، البحث الاجتماعي، دار المعرفة الجامعية، الطبعة الأولى، الإسكندرية.

العطية، فوزية، المدخل إلى دراسة علم النفس الاجتماعي، بغداد، دار الحكمة.

علام، صلاح الدين محمود، تطورات معاصرة في القياس النفسي والتربوي، جامعة الكويت، 1986.

علام، صلاح الدين محمود، القياس والتقويم التربوي والنفسي، القاهرة، دار الفكر العرب، 2000.

علام، صلاح الدين محمود، الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية، القاهرة، دار الفكر العربي، 1995.

عيسوي، عبد الرحمن محمد، دراسات سيكولوجية، منشأة المعارف الإسكندرية.

عيسوي، عبد الرحمن محمد، علم النفس بين النظرية والتطبيق، القاهرة، دار الكتب الجامعية في الإسكندرية.

العيسوي، عبد الرحمن، القياس والتجريب في علم النفس والتربية، الطبعة الأولى، دار النهضة العربية، بيروت.

عاقل، فاخر، علم النفس التربوي، دار العلم للملايين، بيروت، الطبعة الثالثة.

عكاشة، محمود فتحي وآخرون، مقدمة في مناهج البحث التربوي، الطبعة الأولى.

عودة، احمد سليمان، القياس والتقويم في العملية التدريسية، دار الأمل، الأردن، 1998.

الغامدي، سعيد حسن آل عبد الفتاح، مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير عدد بدائل الاستجابة والمرحلة الدراسية، جامعة أم القرى، كلية التربية، المملكة العربية السعودية، رسالة ماجستير غير منشورة، 2003.

الغريب، رمزية، التقويم والقياس النفسي والتربوي، مكتبة إلا نجلو المصرية القاهرة، الطبعة الأولى.

غنيم، سيد محمد، سيكولوجية الشخصية، محدداتها، قياسها، نظرياتها، القاهرة، دار النهضة.

غنيم، سيد محمد، الاختبارات الإسقاطية، القاهرة، دار النهضة العربية.

فاتيحي، محمد، مناهج القياس وأساليب التقييم، مطبعة النجاح الجديدة، الدار البيضاء، 1995.

فإن دالين، مناهج البحث في التربية وعلم النفس، الطبعة الثانية، مكتبة إلا نجلو المصرية، القاهرة.

فراير، دوكلاس هنري، ترجمة إبراهيم المنصور، علم النفس العام، الطبعة الثانية، بغداد

فرج، صفوت، القياس النفسي، الطبعة الأولى، دار الفكر العربي.

فهمي، مصطفى، علم النفس أسسه وتطبيقاته التربوية، القاهرة، مكتبة الخانجي، 1975.

القاضي، يوسف مصطفى وآخرون، الإرشاد النفسي والتوجيه التربوي، دار المريخ، السعودية.

القريوتي، يوسف فريد، مقياس الاتجاهات نحو المتخلفين عقليا، المجلة العربية للعلوم الإنسانية، العدد 29، المجلد الثامن، الكويت.

كرمة، منير جبريل، كيفية بناء جدول المواصفات، برنامج التعليم المفتوح، وكالة الغوث الدولية، فلسطين.

لازارسن، ريتشارد. س، الشخصية، الطبعة الثانية.

لوفيك، ك ولوسون ك، ترجمة إبراهيم البسيوني، لكي نفهم البحث التربوي، الطبعة الأولى، دار المعارف، مصر

مخيمر، صلاح وآخرون، سيكولوجية الشخصية، القاهرة، مكتبة الأنجلو المصرية.

معروف، صبحي عبد اللطيف، الاتجاهات والميول في التربية، منشورات عالم المعرفة

مليكة، لويس كامل، الشخصية وقياسها، الطبعة الأولى، القاهرة، مكتبة النهضة المصرية.

المفتي، كريمة أسعد، دراسة تجريبية لبناء مقياس الميل إلى الدراسة الهندسية، جامعة بغداد، رسالة ماجستير.

منصور طلعت وآخرون، أسس علم النفس العام، القاهرة، مكتبة الأنجلو المصرية.

موحي، محمد آيت، الأهداف التربوية، الطبعة الثالثة، دار الخطابي للطبع والنشر، المغرب العربي، 1988.

النابلسي، محمد احمد، إسقاط الشخصية في اختبار تفهم الموضوع، بيروت، الطبعة الأولى، دار النهضة العربية.

نشواتي، عبد المجيد، علم النفس التربوي، الطبعة الثالثة، دار الفرقان، عمان، الأردن 1987.

النفيعي، عبد الرحمن بن عبد الله بن احمد، تقنين اختبار رافن للمصفوفات المتتابعة المتقدم على طلاب المرحلتين المتوسطة والثانوية بمنطقة مكة المكرمة، جامعة أم القرى، المملكة العربية السعودية، رسالة ماجستير غير منشورة، 2001.

هنا، عطية محمود، الشخصية والصحة النفسية، القاهرة، مكتبة النهضة المصرية.

هنا، عطية محمود، اختبار القيم، دار القيم للنشر والتوزيع، الكويت.

هول، كالفن والآخرون، ترجمة فرج احمد فرج، نظريات الشخصية الهيئة المصرية للتأليف والنشر، القاهرة.

وجيه، إبراهيم وآخرون، البحوث النفسية والتربوية، دار المعارف، القاهرة.

ياسين، عطوف محمود، مدخل في علم النفس الاجتماعي، دار النهار للنشر، بيروت.

ياسين، عطوف محمد، اختبارات الذكاء والقدرات العقلية بين التطرف والاعتدال، الطبعة الأولى، بيروت، دار الأندلس، 1981.

يعقوب، آمال احمد، علم النفس الاجتماعي، جامعة بغداد، دار الحكمة.

يونس، انتصار، السلوك الإنساني، مصر، دار المعارف.

المصادر الأجنبية

Allport G. W. Reading in attitude theory and measurement, newyork, Wiley

Anastasi, A. Psychological testing, newyork, the memillion com,

Angoff, W. H. Scales, Norms and Equivalent scores. inR. L. Thorndik (ED), Educational Measurement, washington, D. C. American council on education.

Berk, R. A. 1980 criterion Referenced Measurement the state of art –Berk, 1986, consumers Guide to setting performance standard on criterion Referenced Test. Review of educational Research. 56,137-172.

Burt, c, Intelligence and social mobility-coombs, theory and method of social measurement in festinger, newDelli, America

Crocker. I, & Algina, J. Introduction to classical & Modern test theory, new york holt. Rinehart & winston.

Down, N. fundamental of measurement London, oxford university press.

Ebel, R. L. Essentials of Educational measurement and ed, Englewood cliffs, n. J.

Edwards, A. L. techniques of attitude scale construction, newyork

Emrick, J, A, An evaluation models for mastery testing. Journale of Educational measurement, 8, 1971.

Encyclopedia of educational Research, 5thed newyork, the free press

Glassier, R.، Introduction technology & the Measurement of learning out comes: some questions Educational Measurement issues & practice, 13, 1994.

Guidance, center catalogue 1984-1985, test books Guidance materials instructional aids, Toronto

Guilford, J. P. Psychometric methods newdelhi, Mccrow- Hillpub. Com. 2ed

Guliksen. H. theory of mental test

Hambleton & Eignor D. R, Competency test development validation and standard setting INR 1980

Hambleton R. K. & Novick M. R. Toward an integration of theory and method for criterion referenced tests, JEM, 10, 1973

Hambleton R. K. on the use of cutt off score with criterion referenced tests in instructional setting JEM, 15, 1978.

Haplin, G < sigmon & haplin, G.، Minimum competency standardset by three

divergent groups of raters using three judgmental procedures implications for validity. Educational & psychological measurement 43, 1983.

Jaeger, R. M. Certification of student competence. IN. linn, educational measurement, newyork, macmillian 1989.

Jensen, A. R. How much we can Boots IQ scholostic achievement, Harvard Educational Review.

Koball, T. R. attitude and Related conception science education, vol. 1,72,No, 1988

Lin dquist, E. F. Educational measurement

Meskauskas, JA Evaluation models for criterion referenced testing, views regarding mastery & standard setting, Review of Educational Research. 46, 1976.

Muller. J. D. Measuring social attitudes, New York

Nedellsky, L.، absolute grading standards of objectives tests, Educational and psychological measurement, 4, 1954.

Nitko m A. j. Distinguishing the many varieties of criterion referenced measurement tests, review of educational research, 50, 1980

Popham & Husenk, T. R. Implication of criterion referenced measurement JEM. 6,1969

Popham, w. j. Criterion Referenced measurement, Englewood, cliffs, N. J. prentice hall.

Sarason G. Irawin, personality nonobjective approach 2n Ed John Willy U. S. A.

Show, marvin E. Scales for measurement of attitude network, McGraw-Hill

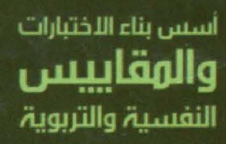
Stevens, S. S.، hand book of Experimental psychology

Vernon, P. E. the structure of human abilities newyork, wiley

Williaw A. Mherens, Irvin J. measurement and evaluation in educational and psychology, New York.


De Bono Center for Teaching Thinking
مركز ديبونو لتعليم التفكير
تفكيرنا... مستقبلنا

أسس بناء الاختبارات
والمقاييس
النفسية والتربوية









traininingcenter

al_debono

مركز ديبونو لتعليم التفكير

هـاتف : 7003 533 6 00962
00962 6 533 7029
فاكس : 7007 533 6 00962
ص.ب : 831 الجبيهة 11941
المملكة الأردنية الهاشمية

E-mail:info@debono.edu.jo
www.debono.edu.jo



9789957454258